Mechanisms of G4C2 derived dipeptide repeat protein toxicity in Drosophila models of C9ALS/FTD

Solomon, Daniel Adam

Awarding institution:
King's College London

Mechanisms of G_4C_2 derived dipeptide repeat protein toxicity in *Drosophila* models of C9ALS/FTD

A thesis submitted to King's College London in fulfilment of the degree of Doctor of Philosophy

(Basic and Clinical Neuroscience)

By

Daniel Adam Solomon, BSc (Hons), MSc

Department of Basic and Clinical Neuroscience

Institute of Psychiatry, Psychology and Neuroscience

King's College London

November 2017

Declaration

I hereby declare that, unless explicitly stated, all of the work presented in this thesis is my own.

Daniel Solomon

October 2017

Acknowledgements

This thesis would not have been possible without the help from numerous people. Firstly I would like to thank my supervisor Dr. Frank Hirth; firstly, for giving me the opportunity to study for a PhD and going out of his way to help me before I started my PhD. I would like to thank Frank for his guidance, always having an open door to talk, his faith in my abilities as a scientist and for always keeping a positive attitude towards the project even when nothing was working and I was naval-gazing as their seemed to be a major advance in C9 every week and we were stuck trying to figure out why our model didn't make much sense. A special thank you to my second supervisor Dr. Jean-Marc Gallo, not only for his scientific input and invaluable support during this PhD but for always also having an open door and being able to chat generally about science and non-science related subjects which I always enjoyed.

One of the highlights of my PhD without doubt has been the people I have met during my time at Kings'. Firstly, I would like to thank Dr. Alan Stepto; we worked together on the same C9 project together for a large proportion of our PhDs. This project was not straightforward and for a really, really, really (reeeeallllllly) long time nothing seemed to work or make sense; however, our shared sense of humour made those times bearable, even enjoyable (dink*)? So, thanks to Alan, even if to this day he has never made me a sandwich. A huge thanks Jessika Bridi, probably the hardest working human being I have ever met for always laughing at all my jokes and dealing with my constant requests to go for lunch from 12pm onwards. Also thank you to Wing Hei Au for her invaluable help and contribution to several of the experiments in this thesis (and the figures); I couldn't have asked for a better and more productive person to work alongside, her approach when it comes to lab work (and expertise with Adobe Illustrator) is incredible. Thanks also to Dr. Yoshi Adachi, Dr. Dani Diaper and Dr. Younbok Lee for all their help, especially Younbok for his enthusiasm and willingness to always answer my questions. Massive thanks to Dr. Claire Troakes and Dr. Tibor Hotrobagyi for all their help and advice with the post mortem tissue work. Thank you to Adel Boudi for his help with the semithin retinal sections. I would also like to show my appreciation to all the BSc and MSc students that had the unfortunate honour of me supervising their projects. Not only for their

contribution to my work but their enthusiasm which always gave me a boost. Thank you to all the great friends I've made, firstly everyone from the 3rd floor back at the IoP main building and secondly everyone who I met when we moved to the Wohl; when the powers that be finally decided to install the sink that was preventing our lab from moving in with everyone else for nearly a year (I'm not joking, and no one even uses the stupid sink. I hate you sink).

I would like to thank Dr. Frank Hirth, Dr. Alan Stepto and Wing Hei Au for their comments and criticisms of this thesis.

A big thank you MNDA for funding my research over the past 4 years, hopefully it wasn't a waste.

Above all I would like to thank my family for all their support. Firstly, my Mum and Dad, for always supporting me and sacrificing so much for my education even if it took a while for the message to get through. Thanks to my brother Jamie and also my dog Ruby; one for their unwavering support and the other for always jumping on me with excitement whenever I had the time to come home. Jamie stop jumping on me it's weird. Above all I would like to thank my Grandparents; my thesis is dedicated to them.

Ok. That's about it. Enjoy reading mates.

Publications arising from this thesis

Oral presentations

A cascade of dipeptide repeat protein and subsequent TDP-43 accumulation causes disease formation in a Drosophila model of C9 ALS/FTD

KCL Neuroscience symposium, *London. UK.* 2016

Daniel Solomon*, Alan Stepto*, Wing Hei Au, Yoshitsugu Adachi, Danielle Diaper, Younbok Lee, Jessika Bridi, Jonah Dearlove, Greta Spinelli, Rachel Hall, Natalie Kirkland, Jean-Marc Gallo, Bradley Smith, Claire Troakes, Richard Parsons, Matthias Soller, Christopher E. Shaw, and Frank Hirth.

Poster presentations

A cascade of dipeptide repeat protein and subsequent TDP-43 accumulation causes disease formation in a Drosophila model of C9 ALS/FTD

KCL Neuroscience symposium, *London. UK.* 2016

Daniel Solomon*, Alan Stepto*, Wing Hei Au, Yoshitsugu Adachi, Danielle Diaper, Younbok Lee, Jessika Bridi, Jonah Dearlove, Greta Spinelli, Rachel Hall, Natalie Kirkland, Jean-Marc Gallo, Bradley Smith, Claire Troakes, Richard Parsons, Matthias Soller, Christopher E. Shaw, and Frank Hirth.

A cascade of dipeptide repeat protein and subsequent TDP-43 accumulation causes disease formation in a Drosophila model of C9 ALS/FTD

4th RNA Metabolism in Neurological Disease, *San Diego. USA.* 2016.

Daniel Solomon*, Alan Stepto*, Wing Hei Au, Yoshitsugu Adachi, Danielle Diaper, Younbok Lee, Jessika Bridi, Jonah Dearlove, Greta Spinelli, Rachel Hall, Natalie Kirkland, Jean-Marc Gallo, Bradley Smith, Claire Troakes, Richard Parsons, Matthias Soller, Christopher E. Shaw, and Frank Hirth.

Modelling the G4C2 repeat expansion in vivo

Origin and evolution of the nervous system, The Royal Society, *London. UK*. 2015

Daniel Solomon*, Alan Stepto*, Wing Hei Au, Yoshitsugu Adachi, Danielle Diaper, Younbok Lee, Jessika Bridi, Jean-Marc Gallo, Bradley Smith, Claire Troakes, Christopher E. Shaw, and Frank Hirth.

Abbreviations

Aβ Amyloid beta

AD Alzheimer's disease

ADARB2 Adenosine Deaminase, RNA-Specific, B2

ALS Amyotrophic lateral sclerosis

ALSbi Amyotrophic lateral sclerosis with behavioural impairment

ALSci Amyotrophic lateral sclerosis with cognitive impairment

ALS-FTD Amyotrophic lateral sclerosis-frontotemporal dementia

ALYREF Aly/REF export factor

AMO Antisense morpholino oligonucleotide

ANOVA Analysis of variance

ASO Anti-sense oligonucleotide

APS Ammonium persulfate

ATG Adenosine-thymine-guanine

ATF4 Activating transcription factor 4

attB Bacterial attachment site

attP Phage attachment site

ATXN2 Ataxin 2

ATXN8 Ataxin 8

ATG13 Autophagy-related 13

ATG101 Autophagy-related 101

AUG Adenosine-uracil-guanine

Aβ Amyloid β

BAC Bacterial artificial chromosome

BCA Bicinchoninic acid

BIP Binding immunoglobulin protein

BSA Bovine serum albumin

bvFTD Behavioural variant frontotemporal dementia

C4G2 (Cytosine)4(guanine)2

C9G4C2 Repeat expansion in C9ORF72

C9L C9ORF72 long isoform

C9ORF72 Chromosome 9 open reading frame 72

C9S C9ORF72 short isoform

CAG Cytosine-adenosine-guanine

CAS Cellular apoptosis susceptibility

CGG Cytosine-(guanine)2

CHAPS 3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate

CHMP2B Charged multivesicular body protein 2B

CHOP CCAAT/Enhancer-binding protein homologous protein

CLEAR Coordinated Lysosomal Expression and Regulation

CNS Central nervous system

CT Computerized tomography

CpG Cytosine-phosphate-guanine

Crm1 Exportin 1

CSF Cerebrospinal fluid

CUG Cytosine-uracil-guanine

DAB 3,3′-Diaminobenzidine

DAPI 4', 6-Diamidino-2-phenylindole

ddH2O Double-distilled water

DENN Differentially expressed in normal and neoplastic cells

DENNL72 DENN-like 72

dFMRP Drosophila Fragile-X mental retardation protein

DM1 Myotonic dystrophy type 1

DM2 Myotonic dystrophy type 2

DMPK Myotonic dystrophy protein kinase

DNA Deoxyribonucleic acid

DN Dystrophic neurites

DPR Dipeptide repeat protein

DTT Dithiothreitol

DSBs DNA double strand breaks

DsRed2 Red fluorescent protein from Discosoma 2

dUTP deoxyuridine triphosphate

EAAT2 Excitatory amino acid transporter 2

EDTA Ethylenediaminetetraacetic acid

EF Elongation factor

eIF Eukaryotic initiation factor

EGFP Enhanced green fluorescent protein

ER Endoplasmic reticulum

EtBr Ethidium Bromide

FALS Familial amyotrophic lateral sclerosis

FG Phenylalanine:glycine

FIP200 FAK family kinase-interacting protein of 200 kDa

FISH Fluorescent in situ hybridization

FFTD Familial frontotemporal dementia

FLCN Folliculin

FNIP1/2 FLCN interacting proteins 1 and 2

FMRP Fragile X mental retardation protein

FMRpolyA Polyalanine

FMRpolyG Polyglycine

FTD Frontotemporal dementia

FTD-MND Frontotemporal dementia-motor neuron dysfunction

FTD-TDP FTD with TDP-43 pathology

FUS Fused in sarcoma

FXTAS Fragile-X tremor/ataxia syndrome

GAPDH Glyceraldehyde 3-phosphate dehydrogenase

GGGGCC G_4C_2 (Guanine)4(cytosine)2

GC Granular component

GCI (oligodendro)glial cytoplasmic inclusion

GDI GDP dissociation inhibitor

GDP Guanine di-phosphate

GEF Guanine exchange factor

GFP Green fluorescent protein

G-quadruplex Guanine-quadruplex

GTP Guanine tri-phosphate

GWA Genome wide association

GWAS Genome wide association study

HCl Hydrochloric acid

HD Huntington's disease

HEK Human embryonic kidney

hnRNP Heterogeneous nuclear ribonucleoprotein

HRP Horseradish peroxidase

HSP Hereditary spastic paraplegia

HSP70 70 kilodalton heat shock proteins

IBB Importin beta binding domain

IDP Intrinsically disordered protein

IFNK Interferon, Kappa

ILF Interleukin Enhancer Binding Factor

IMS Industrial methylated spirit

IPSC Induced pluripotent stem cell

iPSN Induced pluripotent stem cells differentiated into neurons

IRES Internal ribosomal entry site

kb Kilobase

kDa Kilodaltons

KPNB1 Karyopherin Subunit Beta 1

KPNA1 Karyopherin alpha 1 (Importin alpha 5)

KPNA2 Karyopherin alpha 2 (Importin alpha 1)

KPNA3 Karyopherin alpha 3 (Importin alpha 4)

KPNA4 Karyopherin alpha 4 (Importin alpha 4)

KPNA5 Karyopherin alpha 5 (Importin alpha 6)

KPNA6 Karyopherin alpha 6 (Importin alpha 7)

KPNA7 Karyopherin alpha 7 (Importin alpha 8)

L1 First instar larval stage

L3 Third instar larval stage

LCD Low complexity domain

LLPS Liquid-liquid phase separation

LMN Lower motor neuron

LSD Least significant difference

lvPPA Logopenic variant PPA

m^7G Methyl-7-guanosine

MAB414 Anti-Nuclear Pore Complex Proteins antibody

MAPT Microtubule-associated protein tau

Mb Megabase

MBNL Muscleblind

MND Motor neuron disease

MRI Magnetic resonance imaging

mRNA Messenger RNA

Ms-C9orf72 Mouse C9ORF72

Ms-C9orf72-1 Mouse C9ORF72 isoform 1

Ms-C9orf72-2 Mouse C9ORF72 isoform 2

Ms-C9orf72-3 Mouse C9ORF72 isoform 3

NaCl Sodium Chloride

NCI Neuronal cytoplasmic inclusion

NCL Nucleolin

NEB New England Biolabs

NES Nuclear export signal

NF-Kb Nuclear factor kappa-light-chain-enhancer of activated B cells

nfvPPA Nonfluent variant primary progressive aphasia

NGS Normal goat serum

NII Neuronal intranuclear inclusion

NLS Nuclear localisation signal

NLS Nuclei lysis solution

NPC Nuclear pore complex

NPM1 Nucleophosmin 1

Nup Nucleoporin

OBB Odyssey blocking buffer

OPTN Optineurin

p53 Tumour protein p53

p62 Ubiquitin-binding protein p62/Sequestosome 1

PABPc Polyadenylate-binding protein 1

PB Sodium phosphate buffer for immunohistochemistry

PBL Sodium phosphate buffer with lysine hydrochloric acid

P-body Processing body

PBP Progressive bulbar palsy

PBS Phosphate buffered saline

PBST Phosphate buffered saline-Tween

PBT Sodium phosphate buffer with triton X-100

PCR Polymerase chain reaction

PERK Protein kinase RNA-like endoplasmic reticulum kinase

PET Positron emission tomography

PGRN Progranulin

PI Performance index

PIC Pre-initiation complex

PLP Paraformaldehyde fixation solution

PLS Primary lateral sclerosis

PBP Progressive bulbar palsy

PMA Progressive muscular atrophy

PNFA Progressive non-fluent aphasia

Poly(A) Polyadenylation

Poly-GP Poly glycine-proline

Poly-GA Poly glycine-alanine

Poly-GR Poly glycine-arginine

Poly-PA Poly proline-alanine

Poly-PR Poly proline-arginine

PPA Primary progressive aphasia

PPR Poly-pentapeptide repeat

p-Pol-II Phosphor RNA polymerase II

PrLD Prion-like domain

PSD-95 Postsynaptic density protein 95

PY-NLS Proline-tyrosine nuclear localization signal

RAN RAs-related Nuclear protein

RanGAP1 Ran GTPase activating protein 1

RAN translation Repeat-associated non-AUG translation

RBP RNA-binding protein

RCC1 Regulator Of Chromosome Condensation 1

Ref(2)P Refractory to sigma P

RFP Red fluorescent protein

RIPA Radioimmunoprecipitation assay

RNA Ribonucleic acid

RNP Ribonucleoprotein

RO RNA only

ROS Reactive oxygen species

RRM RNA-recognition motif

rRNA Ribosomal RNA

RT Room temperature

SALS Sporadic amyotrophic lateral sclerosis

SC35 Serine/arginine rich splicing factor 2

SCA3 Spinocerebellar ataxia type 3

SCA7 Spinocereballar ataxia type 7

SCA8 Spinocereballar ataxia type 8

SCA10 Spinocereballar ataxia type 10

SCA31 Spinocereballar ataxia type 31

SD Semantic dementia

SDS Sodium dodecyl sulphate

SDS-PAGE Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

SF2 Serine/arginine rich splicing factor 1

SFTD Sporadic frontotemporal dementia

SG Stress granule

shRNA Short hairpin RNA

siRNA Small interfering RNA

SMA Spinal muscular atrophy

SMN1 Survival motor neuron 1

SMRC8 Smith-Magenis Syndrome chromosome region candidate 8

SNP Single nucleotide polymorphism

snRNA Small nuclear RNA

SPECT Single positron emission computerized tomography

SOD1 Superoxide dismutase 1

SQSTM1 Sequestosome 1

SRSF1 Serine And Arginine Rich Splicing Factor 1

SSC Saline sodium citrate

SURF6 Surfeit Locus Protein 6

svPPA Semantic variant primary progressive aphasia

TARBP2 Transactive DNA response RNA-binding protein 2

TARDBP Transactive DNA response DNA-binding protein

TBK1 TANK-Binding Kinase 1

TBPH TAR DNA binding protein homologue

TBS Tris buffered saline

TDP-43 Transactive DNA response DNA-binding protein with molecular weight of 43 kiloDaltons

Tdp-43 Mouse TDP-43

TEMED Tetramethylethylenediamine

TFEB Transcription factor EB

TIA-1 T cell-restricted intracellular antigen-1

TMPyP4 Tetra-(N-methyl-4-pyridyl)porphyrin

TMS Transcranial magnetic stimulation

TNPO1 Transportin 1

UAS Upstream activating sequence

UBQLN2 Ubiquilin-2

UG Uracil-guanine

UMN Upper motor neuron

ULK1 Unc-51-like kinase 1

Unc-119 Uncoordinated 119

UPR Unfolded protein response

UPS Ubiquitin-proteasome system

UTR Untranslated region

UV Ultraviolet

V1 C9ORF72 variant 1

V2 C9ORF72 variant 2

V3 C9ORF72 variant 3

VCP Valosin-containing protein

VNC Ventral nerve cord

WDR41 WD Repeat Domain 41

XPC Xeroderma pigmentosum XPC

Zfp106 Zinc Finger Protein 106

Contents

Abstract

A G_4C_2 hexanucleotide repeat expansion in the gene *C9ORF72* is the most common cause of Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal dementia (FTD). Although intronic, the G_4C_2 repeat is translated in the absence of an ATG start codon through a mechanism known as repeat-associated non-ATG translation resulting in the production of five different dipeptide-repeat proteins (DPRs). These DPRs form inclusions in brains of C9+ ALS and FTD sufferers. Further, C9+ patients are also characterised by cytoplasmic inclusions of the protein TDP-43, whose dysfunction is causally related to ALS and FTD. However, the pathogenic mechanisms underlying C9ALS/FTD and its relation to TDP-43 remain elusive. I have characterised numerous transgenic *Drosophila* to study G_4C_2 and DPR toxicity, and its relation to TDP-43 dysfunction. These *Drosophila* models of C9ALS/FTD are characterised by behavioural deficits that varied in age of onset and severity. The characteristics of motor impairment and subsequent neurodegeneration were dependent on the levels and identity of DPRs produced rather than G_4C_2 RNA. Severe, early onset phenotypes correlated with the presence of poly-GR whereas high levels of poly-GP and poly-GA correlated with late onset phenotypes. Non-G_4C_2 derived poly-GR caused a cytoplasmic accumulation of *Drosophila* TDP-43 (*TBPH*) whilst non-G_4C_2 derived poly-GA sequestered *TBPH* into inclusions. No change in TBPH localisation was seen following G_4C_2 RNA expression only. These alterations were accompanied with changes in the localisation of importins α2 and α3, the *Drosophila* homologues of karyopherins KPNA2 and KPNA4, respectively. However, expression and localisation of nucleocytoplasmic transport components such as RanGAP and nuclear pore complex proteins previously implicated in C9ALS/FTD appeared unperturbed. Further, cytoplasmic mislocalisation of *TBPH* enhanced DPR levels and cytotoxicity. Similar phenotypes were observed for patient frontal cortex in sporadic FTD and C9FTD; with KPNA4 being depleted from the nucleus and overlapping with TDP-43 inclusions. Taken together, these findings establish DPR accumulation as a cause of TDP-43 proteinopathy.

Chapter 1: Introduction

1.1 Motor neuron disease

The motor neuron diseases (MND) are a collection of neurodegenerative disorders characterised by dysfunction of the upper motor neurons (UMNs) of the primary motor cortex and/or dysfunction of the lower motor neurons (LMNs) of the brainstem and spinal cord (figure 1.1). The specific type of motor neurons affected determines the clinical presentation of the disease. Degeneration of the cortical upper motor neurons leads to symptoms such as spasticity, muscle weakness, hyperreflexia and hypertonia. Examples of lower motor neuron dysfunction include atrophy of muscles, fasciculation, muscle weakness, hypotonia, hyporeflexia and cramps. The MNDs can be classified into 6 main subtypes according to the whether the UMNs or LMNs are involved. The UMN diseases include primary lateral sclerosis (PLS) and hereditary spastic paraplegia (HSP). MNDs affecting primarily the LMNs include progressive muscular atrophy (PMA), progressive bulbar palsy (PBP) and spinal muscular atrophy (SMA). And finally the most common form of MND is amyotrophic lateral sclerosis (ALS) which presents with a combined UMN and LMN involvement.

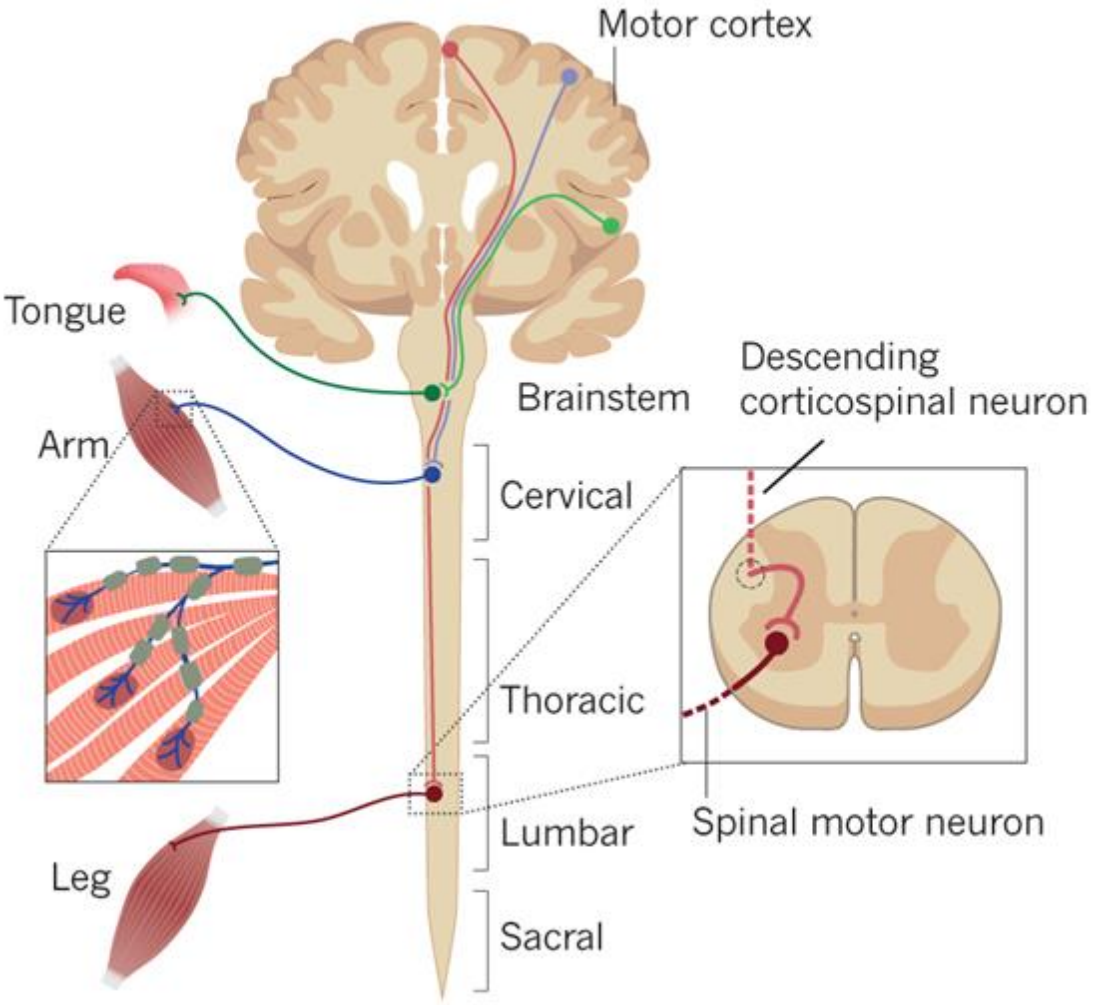

Figure 1.1 | Upper and lower motor neurons in the brain and spinal cord. Upper motor neurons (UMNs) originate from the motor cortex of the brain and synapse with lower motor neurons (LMNs) either in the brain stem (corticobulbar tract; green) or spinal cord (anterior and lateral corticospinal tracts; blue & red). LMNs originating from the brainstem synapse with the muscles of the face, head and neck. LMNs of the spinal cord synapse innervate muscles of the digits, limbs and trunk of the body. Adapted from Taylor et al. (2016).

1.1.1 Primary Lateral Sclerosis

PLS is a rare form of MND in which patients present with upper motor neuron symptoms only (Swinnen and Robberecht, 2014). PLS accounts for around 5% of MND cases (D'amico et al., 2013). Disease progression is slow compared to other forms of MND with sufferers retaining core LMN features and living longer than patients with ALS (Pringle et al., 1992; Tartaglia et al., 2007). PLS is usually associated with focal cortical atrophy in precentral regions (Kuipers-Upmeijer et al., 2001). Despite being a UMN disease, LMN dysfunction has been shown to develop over time (Le Forestier et al., 2001). Further, many of the pathogenic hallmarks of ALS are seen in PLS, such as ubiquitinated inclusions and bubina bodies (Tartaglia et al., 2007). This had led to the suggestion that PLS may be a rare slowly progressive form of ALS.

1.1.2 Hereditary Spastic Paraplegia

The HSPs are a clinically and genetically diverse group of diseases typified by a progressive weakness and spasticity of the lower limbs (Lo Giudice et al., 2014). Neuronal death is typically rare in HSP; rather HSP is a distal axonopathy that causes degeneration of the long corticospinal tract axons (Blackstone et al., 2011). HSP can be divided into pure and complicated forms. Pure forms of the disease present only with spasticity of the lower extremities whereas complicated forms of the disease have additional clinical features including cognitive dysfunction, epilepsy, optic atrophy and skeletal abnormalities (Lo Giudice et al., 2014; Blackstone et al., 2011). HSP has a prevalence of around 18 per 100,000 (Blackstone et al., 2011) with pure forms of the disease being more common in Northern Europe, North America and Japan and complicated forms seen more frequently in Mediterranean populations (Lo Giudice et al., 2014).

1.1.3 Progressive Muscular Atrophy

Patients with PMA present with lower motor neuron symptoms exclusively (Kim et al., 2009). 2.5-11% of patients with MND are classified as having PMA (Kim et al., 2009). PMA patients often show at autopsy corticospinal tract degeneration with

surviving motor neurons presenting with ubiquitinated inclusions as seen in ALS (Ince et al., 2003). However, around 20% of PMA patients show UMN symptoms (Kim et al., 2009); additionally in a transcranial magnetic stimulation (TMS) study, around 35% PMA patients were found to have UMN dysfunction (Floyd et al., 2009). Even though PMA patients live longer than classic ALS sufferers; risk factors associated with reduced life-span in PMA are the same for ALS (Kim et al., 2009). Taking these observations together has led some to conclude that PMA may represent a rare form of ALS rather than being a separate disease (Kim et al., 2009).

1.1.4 Progressive bulbar palsy

PBP involves the selective degeneration of motor neurons of the lower brain stem and can present with or without corticobulbar tract complications (Karam et al., 2009). Difficulties in speech and swallowing occur early in the disease and are rapidly followed by dysfunction of the limbs, respiratory system and muscles (Talbot, 2009). PBP is more common in women than men with age of onset typically being over 65 years of age (Talbot, 2009). Some PBP patients present with UMN symptoms, however, independently of this, nearly all PBP patients will go on to develop ALS (Karam et al., 2009). Hence some consider PBP to represent bulbar onset ALS instead of a distinct disease (Swinnen and Robberecht, 2014).

1.1.5 Spinal muscular atrophy

SMA is the term for a collective group of diseases that affect the motor neurons in the brainstem and spinal cord (Faravelli et al., 2015). The different types of SMA are associated with different genetic mutations and clinical presentations (Arnold et al., 2015). The most common form of SMA is autosomal recessive proximal SMA or 5q-SMA that accounts for 95% of cases (Arnold et al., 2015). 5q-SMA is caused either by a homozygous deletion or mutation in the gene *Survival Motor Neuron 1* (*SMN1*) (Faravelli et al., 2015; Arnold et al., 2015). 5q-SMA is the most common cause of infant death and affects around 1:10,000 births (Prior et al 2010). The 5q-SMA disease spectrum is divided into 5 different types (type 0 – 4) depending on age of onset, disease severity and age of death (Faravelli et al., 2015). The most common form, type 1, usually occurs at around 6 months of age (Arnold et al., 2015), but 5q-

SMA can present as early as *in utero* (type 0) or as late as the third decade of life (type 4) (Faravelli et al., 2015). Typically, all types of 5q-SMA present with muscle atrophy and weakness at different severities usually starting in the lower limbs (Faravelli et al., 2015).

1.1.6 Amyotrophic lateral sclerosis

ALS is the most common form of MND (Lattante et al., 2015). ALS has a European incidence rate of around 2.16 per 100,000, with the disease being slightly more prevalent in men than women (3 per 100,000 and 2.4 per 100,000 respectively) (Logroscino et al., 2010). ALS is characterised by the degeneration of cortical UMNs together with degeneration of bulbar and spinal LMNs (figure 1.2) (Swinnen and Robberecht, 2014). Disease onset is typically after the age of 40, usually in fifth or sixth decade of life, although there are rare juvenile forms of ALS beginning before the age of 25 (Robberecht and Phillips, 2013; Swinnen and Robberecht, 2014). Around 10% if patients have a family history of the disease where the cause of most, if not all, is through the inheritance of ALS causing genetic mutations (Swinnen and Robberecht, 2014). Such patients are categorised as having familial ALS (FALS). In the remaining 90% of patients no familial history of the disease is present hence are recorded as having sporadic ALS (SALS). The causes of SALS are unknown but it is likely to have both a genetic and environmental component (Swinnen and Robberecht, 2014).

ALS is a heterogeneous disease both genetically and phenotypically (Swinnen and Robberecht, 2014). Patients show significant variability in the precise area of symptom onset (Kiernan et al., 2011). The most common presentation is spinal-onset ALS which accounts for around 70% of cases (Kiernan et al., 2011). Spinal-onset ALS begins with an asymmetric limb weakness (Swinnen and Robberecht, 2014). Patients show classic signs of UMN dysfunction such as hypertonia and hyperreflexia and LMN dysfunction including muscle weakness and atrophy (figure 1.2) (Swinnen and Robberecht, 2014). 25% of patients have a bulbar-onset form of ALS (Kiernan et al., 2011) which typically commences with a weakness in the bulbar muscles accompanied by difficulties in speech and swallowing in addition to wasting of the tongue (Swinnen and Robberecht, 2014). 5% of ALS patients have a respiratory-onset of the disease (Kiernan et al., 2011) involving severe breathing

difficulties that requires rapid respiratory intervention (Gautier et al., 2009). Most patients die within 3-5 years of diagnosis although there is significant variability in disease duration with some patients living for decades (Robberecht and Phillips, 2013). Patients with bulbar-onset ALS have a much shorter survival (2 years mean survival with a 10 year survival rate of 3.4%) compared to those with spinal-onset (2.6 years mean survival with 10 year survival rate of 13%) (Chiò et al., 2011). The prognosis for those with respiratory-onset ALS is bleak, with a mean survival of only 1.4 years and no one surviving up to 10 years (Chiò et al., 2011). Interestingly a younger the age of onset correlates with a longer survival (Pupillo et al., 2014).

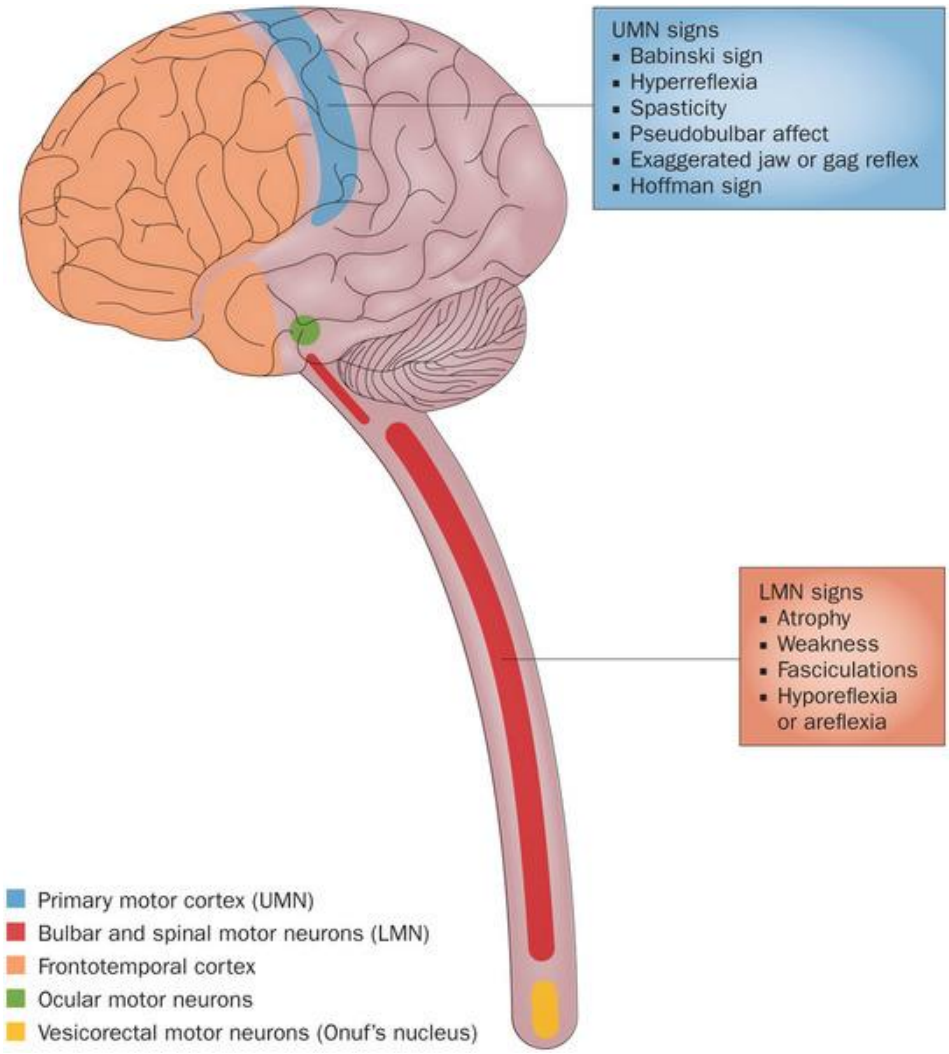

Figure 1.2 | Neurons affected in ALS. ALS involves both the upper motor neurons (UMNs) in the primary motor cortex (blue) and bulbospinal lower motor neurons (LMNs) (red). UMN loss leads to symptoms such as hyperreflexia and spasticity whilst LMN loss leads to muscle atrophy and weakness. ALS also frequently involves neurons in the frontotemporal cortex (orange). The ocular (green) and vesicorectal (yellow) motor neurons are rarely involved. From Swinnen and Robberecht, (2014).

1.2 Frontotemporal Dementia

Frontotemporal dementia (FTD) is a group of clinically and pathologically distinct neurodegenerative diseases characterised by the selective loss of neurons in the frontal and temporal cortices (figure 1.3) (Warren et al., 2013). The prevalence of FTD varies, ranging from 2.7 per 100,000 to 15.1 per 100,000 in different epidemiological studies (Rabinovici and Miller, 2010). A higher number of FTD cases are recorded as familial compared to ALS (Turner et al., 2017); up to 30% of FTD cases are thought to be familial with a known genetic cause (FFTD) whilst the remaining 70% are sporadic (SFTD) (Turner et al., 2017). FTD typically occurs in the sixth decade of life, although can present as early as the third (Rabinovici and Miller, 2010). Clinically FTD patients undergo a progressive decline in behavioural and/or language faculties. Based on the exact clinical presentation of the disease, FTD can be subdivided into three syndromes; behavioural variant frontotemporal dementia (bvFTD), semantic dementia (SD) and progressive non-fluent aphasia (PNFA) (Warren et al., 2013).

The bvFTD subtype of FTD is the most common of the FTD clinical syndromes (Rohrer et al., 2015). Patients with bvFTD show a loss of interpersonal and executive skills that is accompanied by emotional blunting and the appearance of unusual behaviours such as social withdrawal, confrontation seeking, personality changes, overeating, a loss of personal hygiene, obsessions, rituals and even sociopathic behaviours (Rohrer et al., 2015, Warren et al., 2013, Rabinovici and Miller, 2010). The cognitive decline in bvFTD is more subtle, with patients most commonly showing deficits in executive dysfunctions including attentiveness, judgement, planning and organisation (Rabinovici and Miller, 2010). This subtype of FTD is associated with pronounced frontal, insular and anterior cingulate atrophy (figure 1.3) (Meeter et al., 2017).

The SD form of FTD involves the continuous deterioration of semantic memory required for general knowledge about the world accumulated from life experiences (Warren et al., 2013). It accounts for around 20 – 25% of diagnosed FTD cases (Gosh and Lippa, 2015). SD predominantly involves language problems, in particular a difficulty understanding the meaning of words, objects and concepts (Warren et al., 2013, Rabinovici and Miller, 2010). Speech is fluent but patients are unable to

retrieve words, show an impaired knowledge of familiar objects and are commonly unaware of their lack of comprehension (Gosh and Lippa, 2015). Structural magnetic resonance imaging (MRI) of SD shows a typical anteroinferior temporal lobe cortical atrophy which is asymmetric and frequently worst in the left hemisphere (figure 1.3) (Gosh and Lippa, 2015, Warren et al., 2013; Meeter et al., 2017).

The final subtype of FTD is PNFA, presenting in around 25% of FTD patients (Gosh and Lippa, 2015). Both the PFNA and SD forms of FTD can be grouped together into a broader language syndrome known as primary progressive aphasia (PPA) (Gosh and Lippa, 2015). PPA is classified into 3 clinical variants dependent on the precise speech and language deficits observed (Gosh and Lippa., 2015). The first variant is PNFA, also known as non-fluent variant PPA (nfvPPA) (Rohrer et al., 2015). PNFA sufferers show severe language deficits characterised by deterioration in language output and slow, non-fluent speech (Warren et al., 2013). Patients present with agrammatism, speech apraxia and a difficulty in understanding complicated sentences (Gosh and Lippa., 2015, Warren et al., 2013). The second variant is SD, which can also be referred to as svPPA (Rohrer et al., 2015); SD involves loss of semantic memory as described previously (Gosh and Lippa, 2015, Warren et al., 2013). The final form of PPA is termed logopenic variant PPA (lvPPA) or logopenic aphasia (Rohrer et al., 2015); those diagnosed with lvPPA have grammatically correct speech, but this speech is hesitant and marked by constant word-finding pauses and difficulties in repetition (Gosh and Lippa., 2015, Warren et al., 2013). Brain atrophy varies between PFNA patients (Warren et al., 2013) but left-sided frontal and insular atrophy is often reported (figure 1.3) (Meeter et al., 2017)

FTD is a progressive disease; however, survival is variable across the FTD spectrum (Warren et al., 2013). Mean survival is around 6-12 years after diagnosis (Kansal et al., 2016). A recent meta-analysis of 27 FTD survival studies revealed all 3 subtypes of FTD (bvFTD, SD and PNFA) had similar survival rates with mean survival being longest in bvFTD and PNFA (8 years) and median survival being greatest for SD (12 years) (Kansal et al., 2016). Heritability varies between the different FTD syndromes; bvFTD has a stronger familial basis compared to both SD and PNFA, with the majority of SD cases being sporadic (Rohrer et al., 2009).

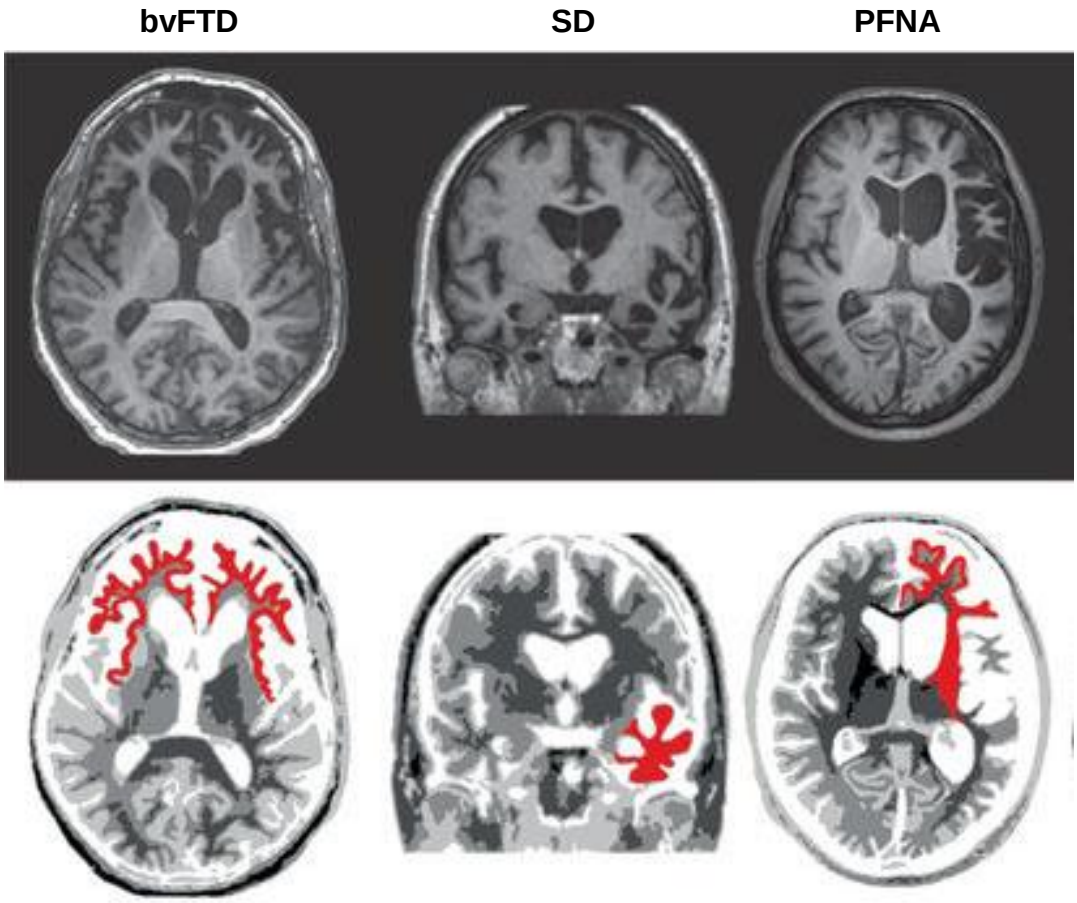

Figure 1.3 | Grey matter atrophy in FTD variants. The behavioural variant subtype of FTD (bvFTD) is associated with a pronounced atrophy of the frontal, insular and anterior cingulate cortices. Semantic dementia (SD) typically involves asymmetrical left-sided temporal atrophy. Progressive non-fluent aphasia (PFNA) patients exhibit left-sided frontal and insular atrophy. Adapted from Meeter et al. (2017).

1.3 The ALS-FTD disease spectrum

Over the past several decades based on clinical, pathological and genetic observations there has been recognition of a significant overlap between ALS and other neurodegenerative diseases (Taylor et al., 2016). Cognitive impairment in seen a significant proportion of ALS patients and motor symptoms resembling ALS are reported in numerous cases of FTD (Taylor et al., 2016). Indeed, the two diseases have been brought together at the clinical, pathological and genetic level (Swinnen and Robberecht, 2014; Ling et al., 2013). ALS and FTD are now thought to lie on a clinical continuum with pure ALS and FTD at the polar ends of this spectrum and a combined syndrome, ALS-FTD, in the centre (figure 1.4) (Swinnen and Robberecht, 2014; Ling et al., 2013).

Figure 1.4 | Phenotypic variability across the ALS-FTD spectrum. ALS and FTD lie on a disease spectrum with ALS on one end and FTD the other. This includes patients with pure ALS, those who have ALS with cognitive or behavioural impairment (ALSci or ALSbi) and those ALS patients who meet the diagnosis for FTD (ALS-FTD). Some FTD patients do not have sufficient motor neuron involvement for an ALS diagnosis so are diagnosed as FTD-MND. Adapted from Swinnen & Robberecht, (2014).

1.3.1 Converging clinical features of ALS and FTD

Around 20% of ALS patients meet the clinical criteria for an accompanying FTD diagnosis, with cognitive dysfunction being reported in almost half of all ALS cases (Taylor et al., 2016). The most common cognitive symptoms experienced in ALS are deficits in verbal fluency, visual memory, attention and working memory (Strong and Yang, 2011; Phukan et al., 2007). In addition to these cognitive deficits, behavioural impairments are also prevalent in patients; typical abnormalities include apathy, disinhibition and the development of delusions and stereotypical behaviours (Swinnen and Robberecht, 2014; Strong and Yang, 2011). ALS patients suffering from mild abnormalities in behaviour are diagnosed as ALS with behavioural

impairment (ALSbi) and those with mild cognitive and language impairment are categorized as having ALS with cognitive impairment (ALSci) (Swinnen and Robberecht, 2014). MRI and computerized tomography (CT) scans reveal ALS patients with these frontotemporal syndromes have degeneration of neurons in the frontotemporal cortex (Strong and Yang, 2011). Both single positron emission computerized tomography (SPECT) and positron emission tomography (PET) imaging techniques have shown deficits in verbal fluency correlate with reduced cerebral blood flow to the anterior cingulate gyrus and regions of both the frontal and temporal cortices (Strong and Yang, 2011).

Motor symptoms are estimated to be found in anywhere between 5-10% of FTD cases (Seelar et al., 2007) to as high as 50% of cases (Swinnen and Robberecht, 2014). Patients who do not have severe enough motor neuron involvement for ALS are categorized as having FTD-MND (Swinnen and Robberecht, 2014). All subtypes of FTD can come concomitant with MND; however, it most commonly occurs with the bvFTD form and more rarely with the language variants of the disease (Seelar at al., 2011). Depending on the patient, MND may occur early or late during FTD progression; with weakness and atrophy of the muscles being the most common symptoms reported, particularly in the arms and tongue (Seelar et al., 2011). Those suffering from ALS-FTD tend to have a uniquely rapid onset of dementia that usually comes before the motor symptoms (Lillo and Hodges, 2009). ALS-FTD patients present with a mixture of language and behavioural deficits (Lillio and Hodges, 2009) with these cognitive deficits usually appearing before signs of obvious motor impairment (Swinnen and Robberecht, 2014). Patients with ALS-FTD have on average worse survival than those with pure ALS or FTD, with a survival average of 2.4 years after diagnosis (Olney et al., 2005, Kansal et al., 2016). In addition to the strong clinical overlap between ALS and FTD, the strongest evidence that these two diseases are two clinical manifestations of the same pathological process comes from research into the pathology and genetics underlying these two disorders (Ling et al., 2013).

1.3.2 Shared pathological features of ALS and FTD

Both ALS and FTD can be categorised into different subtypes in accordance with the histopathology shown in the brain at post-mortem analysis; with both diseases characterised by the presence of proteinaceous inclusions (figure 1.5). Some of the misaccumulated proteins are found only in ALS but not FTD and vice versa. For example, inclusions of superoxide dismutase 1 (SOD1) are found in 2% of ALS cases but not in any FTD patients (Ling et al., 2013). SOD1 is a cytoplasmic and mitochondrial antioxidant enzyme that converts superoxide into molecular oxygen and hydrogen peroxide (Bunton-Stasyshyn et al., 2015). For FTD, 45% of patients harbour pathological protein deposits of microtubule-associated protein tau (MAPT), such inclusions are not found in ALS post-mortem brains (Ling et al., 2013). Tau promotes microtubule assembly and stability; in FTD an abnormal intracellular accumulation of hyperphosphorylated tau is observed (Mackenzie and Neumann, 2016)

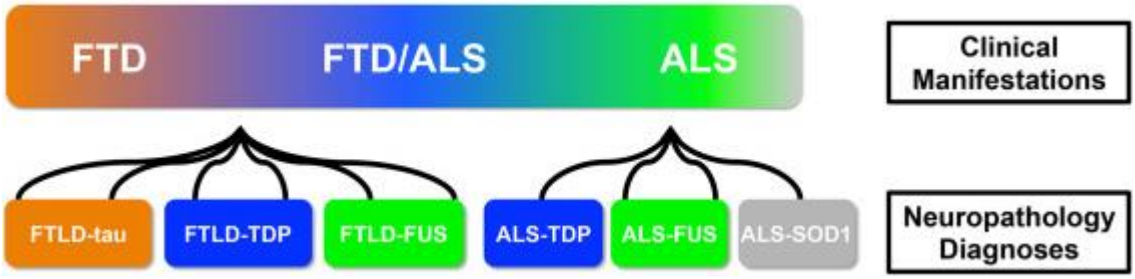

Figure 1.5 | The major neuropathological subtypes in the ALS and FTD disease spectrum. FTD and ALS patients can be classified according to the protein aggregates found in neurons at post-mortem. Some protein deposits are only associated with FTD such as Tau (FTLD-tau); whilst others are only found in ALS such as SOD1 (ALS-SOD1). Other proteins form inclusions in both diseases, such as TDP-43 (FTLD-TDP and ALS-TDP,) and FUS (ALS-FUS and FTLD-FUS). Adapted from Lee and Huang, (2017).

However, several proteins have been found to form protein aggregates in both diseases (figure 1.5). Deposits of the RNA binding protein Fused in Sarcoma (FUS) are found in <1% of ALS and 9% of FTD cases respectively (Ling et al., 2013) (figure 1.6). The major pathological breakthrough linking these two diseases was the discovery that Transactive DNA response (TAR) DNA-binding protein with a molecular weight of 43kDa (TDP-43) is main component of the ubiquitated inclusions

found in both sporadic ALS and the most common pathological form of FTD (Neumann et al., 2006). TDP-43 proteinaceous inclusions are found in 97% of all ALS cases and 45% of FTD brains (Ling et al., 2013) (figure 1.6). TDP-43 is an RNA binding protein with large number of roles in RNA processing (Ling et al., 2013). TDP-43 shuttles between the nucleus and cytoplasm but is predominantly found to be nuclear in its steady state (Ling et al., 2013). In ALS and FTD TDP-43 is cleared from the nucleus forming ubiquitinated and hyperphosphorylated inclusions that aggregate in the cytoplasm (Neumann et al., 2006; Arai et al., 2006). Pathological TDP-43 aggregates consist of full length TDP-43 in addition to N-terminal and C-terminal fragments of the protein (Scotter et al., 2015; Xiao et al., 2015b). The loss of TDP-43 from the nucleus suggests a loss of nuclear TDP-43 function as a central pathogenic mechanism; whilst the presence of cytoplasmic TDP-43 aggregates indicates a toxic gain of function as a potential disease mechanism; indeed, functional studies in animal models support both of these possibilities (Lee et al., 2011; Ling et al., 2013). It is likely both a toxic loss of TDP-43 function and a cytoplasmic gain of TDP-43 function play a role in TDP-43 mediated neurodegeneration (Ling et al., 2013). Neurodegenerative diseases with major TDP-43 deposition are termed TDP-43 proteinopathies (Scotter et al., 2015); in such diseases the TDP-43 pathology correlates with neuronal loss and in ALS the regional spread of TDP-43 proteinopathy from the spinal cord and motor neurons to cortical areas can used to stage the progression of the disease (Scotter et al., 2015). Hence understanding TDP-43 dysfunction, and risk factors underpinning the development of TDP-43 proteinopathy - is key to understanding neurodegeneration in these diseases and for the design of therapeutic interventions.

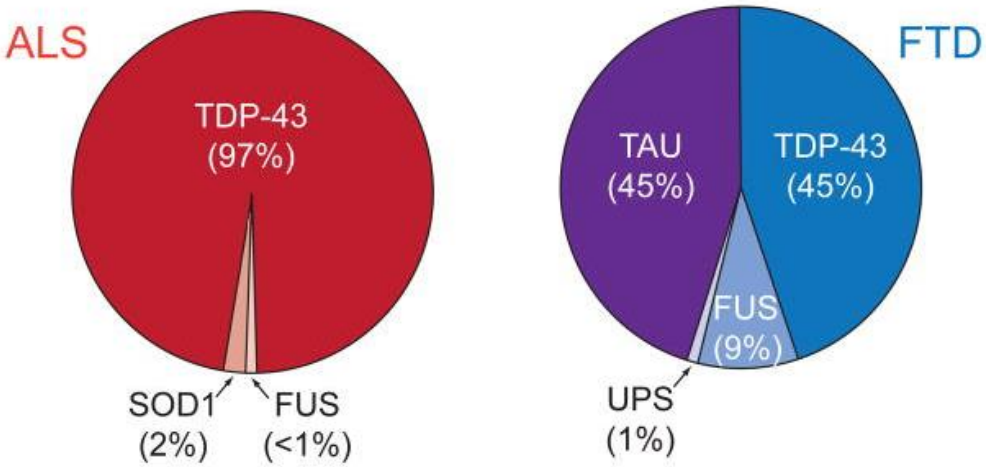

Figure 1.6 | Pathological inclusions in ALS and FTD. Pathological protein inclusions seen in ALS and FTD. The major misaccumlated protein in ALS is TDP-43, being observed in 97% of ALS cases. TDP-43 pathology is also observed in a significant proportion of FTD cases (45%). FUS inclusions are also observed in both ALS (<1%) and FTD (9%). Hence aggregates of TDP-43 and FUS demonstrate the pathological overlap between ALS and FTD. SOD1 and Tau inclusions are seen only in ALS (2%) and FTD (45%) respectively. Adapted from Ling et al. (2013).

1.3.3 Genetics of ALS and FTD

A direct molecular link between ALS and FTD has come from the identification of genes which when mutated cause both diseases (Ling et al., 2013) (figure 1.7). The most prominent example of this is the GGGGCC (G_4C_2) hexanucleotide repeat expansion in the non-coding region of *Chromosome 9 open reading frame 72* (*C9ORF72*) - the most common single genetic abnormality for both FALS and FFTD (Haeusler et al., 2016). The number of genes known to be associated with both diseases has greatly expanded over recent years with the development of exome sequencing - which enables only the protein coding regions of the human genome, where the majority of human disease-causing mutations are found, to be sequenced.

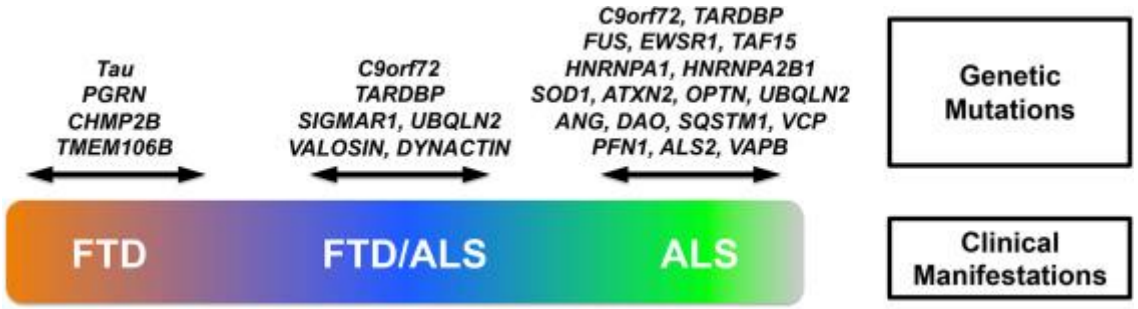

Figure 1.7 | The major genes associated with the ALS and FTD disease spectrum. Both FTD-ALS and ALS have similar genetic mutations, in particular the G_4C_2 repeat expansion in *C9ORF72*. Some genes when mutated cause FTD alone; such as *Tau*. Other genes when mutated cause pure ALS e.g. *SOD1*. Adapted from Lee and Huang, (2017).

Mutations in genes such as *SOD1* cause pure ALS (figure 1.7) whereas mutations in other genes are associated only with FTD including *MAPT* (figure 1.7) (Hardy and Rogaeva, 2014). These mutations give rise to distinct disorders (ALS and FTD respectively) and lead to disease through pathways not involved in other forms of ALS, FTD and ALS/FTD (Hardy and Rogaeva, 2014). A closer look at the functions of the genes which cause ALS alone demonstrates such genes are typically involved in RNA/DNA metabolism; such as *TAR DNA binding protein* (*TARDBP*) and *Fused in Sarcoma* (*FUS*) (Hardy and Rogaeva., 2014). This indicates motor neurons have specific requirements for RNA processing and a disruption in this biological function is particularly detrimental to their function (Hardy and Rogaeva, 2014). Genes which

cause ALS-FTD or FTD alone seem to have functions in the same biological pathways (figure 1.7). These genes are typically associated with protein degradation and a damaged autophagy/lysosomal pathway; examples include mutations in *C9ORF72*, *valosin containing protein* (*VCP*), *sequestosome 1* (*SQSTM1*), *optineurin* (*OPTN*), *ubiquilin-2* (*UBQLN2*), *charged multivesicular body protein 2B* (*CHMP2B*) and *progranulin* (*PGRN*). This suggests FTD and the ALS/FTD disease continuum is associated with inhibited protein degradation and defective autophagy (Hardy and Rogaeva, 2014). Hardy and Rogaeva, (2014) argue the function of the genes associated with ALS and FTD points to common disease mechanisms detrimental to susceptible neuronal populations; i.e. perturbed RNA metabolism in motor neurons and disrupted protein clearance in both motor and cortical neurons (figure 1.8).

Gene official symbol	Gene name	Location	Clinical presentation	Possible pathways / pathological biological processes
MENDELIAN GENES				
MAPT	Microtubule-associated protein tau	17q21.1	FTD	Toxic aggregation (defect in neuronal cytoskeleton)
GRN	Granulin	17q21.32	FTD*	Autophagy; lysosomal pathway; inflammation
TREM2	Triggering receptor expressed on myeloid cells 2	6p21.1	FTD	Inflammation
CHMP2B	Charged multivesicular body protein 2B	3p11.2	FTD	Autophagy; lysosomal pathway
C9orf72	Chromosome 9 open reading frame 72	9p21.2	FTD, ALS	Toxic RNA or repeat dipeptides aggregation
SQSTM1	Sequestosome 1	5q35	FTD, ALS*	Autophagy
UBQLN2	Ubiquilin 2	Xp11.21	FTD, ALS*	Autophagy
VCP	Valosin-containing protein	9p13.3	FTD, ALS*	Autophagy
OPTN	Optineurin	10p13	FTD/ALS*	Autophagy
SOD1	Superoxide dismutase 1, soluble	21q22.11	ALS	Toxic aggregation; free radical scavenger enzyme
FUS	FUS RNA binding protein	16p11.2	ALS	DNA/RNA metabolism
TARDBP	TAR DNA binding protein	1p36.22	ALS	DNA/RNA metabolism
CHCHD10	Coiled-coil-helix-coiled-coil-helix domain containing 10	22q11.23	FTD, ALS	Mitochondrial function
ALS2	Amyotrophic lateral sclerosis 2 (juvenile)	2q33.1	ALS*	Modulator for endosomal dynamics
SPG11	Spastic paraplegia 11 (autosomal recessive)	15q14	ALS*	DNA damage repair
SETX	Senataxin	9q34.13	ALS*	DNA/RNA processing
MATR3	Matrin 3	5q31.2	ALS*	DNA/RNA metabolism
ANG	Angiogenin, ribonuclease, RNase A family, 5	14q11.1-q11.2	ALS	Blood vessel formation
VAPB	VAMP (vesicle-associated membrane protein)-associated protein B and C	20q13.33	ALS*	Vesicle trafficking
PFN1	Profilin 1	17p13.3	ALS	Actin dynamics
TAF15	TAF15 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 68kDa	17q11.1-q11.2	ALS	RNA metabolism
HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1	12q13.1	ALS, FTD*	RNA metabolism; direct interaction with TDP-43
HNRNPA2B1	Heterogeneous nuclear ribonucleoprotein A2/B1	7p15	ALS, FTD*	RNA metabolism; direct interaction with TDP-43
ERBB4	Erb-b2 receptor tyrosine kinase 4	2q33.3-q34	ALS	Dysregulation of the neuregulin-ErbB4 pathway
ARHGEF28	Rho guanine nucleotide exchange factor (GEF) 28	5q13.2	ALS	Interaction with low-molecular-weight neurofilament mRNA
DAO	D-amino-acid oxidase	12q24	ALS	
GLE1	GLE1 RNA export mediator	9q34.11	ALS	RNA metabolism
SIGMAR1	Sigma non-opioid intracellular receptor 1	9p13.3	ALS, FTD	Endoplasmic reticulum lipid rafts
ERLIN2	ER lipid raft associated 2	8p11.2	ALS*	Endoplasmic reticulum lipid rafts
PNPLA6	Patatin-like phospholipase domain containing 6	19p13.2	ALS*	Neurite outgrowth and process elongation
PRKAR1B	Protein kinase, cAMP-dependent, regulatory, type I, beta	7p22	FTD*	Regulation of metabolism, ion transport, and gene transcription
DCTN1	Dynactin 1	2p13	ALS*	Vesicle trafficking

Figure 1.8 | Genetics of ALS and FTD*. Genes which lead to ALS alone are strongly associated with RNA/DNA metabolism, such as *TARDBP* and *FUS*. Genes which cause FTD alone or both ALS and FTD are associated with autophagic processing. Adapted from Guerreiro et al., (2015). *****Table does not include the recent identification of mutations in *TUB4A* in ALS (cytoskeleton) (Smith et al., 2014); *TBK1* in ALS and FTD (autophagy) (Cirulli et al., 2015; Freischmidt et al., 2015); *NEK1* in ALS and FTD (cell cycle regulation) (Cirulli et al., 2015); *CCNF* (E3 ubiquitin-protein ligase complex and cell cycle regulation) (Williams et al., 2016); *ANXA11* in ALS (Vesicular protein trafficking) (Smith et al., 2017); and *TIA1* in ALS and FTD (RNA metabolism) (Mackenzie et al., 2017).

1.4 Chromosome 9 open reading frame 72 (*C9ORF72*)

A linkage of a 9.8 megabase (Mb) region located between chromosome 9p13.3-p21.3 was observed within European ALS and FTD families (Morita et al., 2006; Vance et al., 2006). Linkage analysis managed to narrow the linkage region to 3.7 Mbs but did not identify any disease causing mutations (Boxer et al., 2011; Pearson et al., 2011). Complimentary genome-wide association studies (GWAS) also identified the same region in a Dutch FALS cohort (van Es et al., 2009) and in 47.3% of FALS patients and nearly 25% of ALS patients in a Finnish cohort (Laaksovirta et al., 2010). A larger GWAS study in SALS patients from seven different countries again reported an association with the same region of the chromosome (Shatunov et al., 2010). GWAS studies also identified SFTD with TDP-43 pathology associated with this locus (Van Deerlin et al., 2010). These GWAS studies identified the same single nucleotide polymorphism (SNP), rs3849942, on the chromosome 9p locus (Laaksovirta et al., 2010; Shatunov et al., 2010). The disease associated SNPs narrowed the associated signal on chromosome 9p to a block of linkage disequilibrium ~106.5Kb in size which included three genes; *MOBKL2B*, *INFK* and *C9ORF72* (Shatunov et al., 2010).

The disease causing mutation on chromosome 9p-linked FTD-ALS locus was discovered in 2011 by DeJesus-Hernandez et al. (2011) and Renton et al. (2011). The mutation was discovered to be a massive hexanucleotide repeat expansion of the sequence G_4C_2 within intron 1 of *C9ORF72*. The G_4C_2 repeat number in healthy individuals ranged from 2-23 repeats whereas repeat number in affected individuals was expanded into the hundreds to thousands (DeJesus-Hernandez et al., 2011; Renton et al., 2011). The difficulty in identifying the mutation previously owed to the GC rich nature of the repeats making standard PCR sequencing techniques inadequate. The DeJesus-Hernandez et al. (2011) and Renton et al. (2011) studies got around this problem by utilizing repeat-primed PCR and next-generation sequencing to identify to the presence of the G_4C_2 repeats. This repeat expansion in *C9ORF72* is the most common single genetic abnormality in FALS and FFTD (Haeusler et al., 2016). Across different studies conducted in several populations the expansion is has been shown to account for concomitant FALS-FTD (10-87.5%), SALS-FTD (6.3%-22.2%), FALS (3-46%), SALS (0.4%-21%), FFTD (3-48%) and

SFTD (2-23%) (Cruts et al., 2013). The repeat expansion is particularly prevalent in European and North American populations (Cruts et al., 2013). Majounie et al. (2012b) found in a large cohort of 4448 patients diagnosed with ALS and 1425 patients diagnosed with FTD that the most frequent cause of sporadic ALS and sporadic FTD identified thus far was the expansion, accounting in total for around 5·0–7·0% of cases in white Europeans, Americans and Australians. Particularly high frequencies of the expansion are seen in the Scandinavian countries Finland, Sweden and Denmark with the expansion being rarer in Asian populations; suggesting a northern European founder (Cruts et al., 2013).

The *C9ORF72* gene consists of 12 exons - 10 coding exons (exons 2 – 11) and 2 non- coding exons (exons 1a and 1b) (figure 1.9). (Stepto et al., 2014; Haeusler et al., 2016). Based on intron-exon structure and differential transcriptional start and termination sites the *C9ORF72* gene is alternatively spliced to produce 3 mRNA transcripts in which the non-coding exons 1a and 1b are differentially incorporated into the pre-mRNA. The three transcripts are V1 (NM_145005.5), V2 (NM_0183525.3) and V3 (NM_001256054.1) (Stepto et al., 2014); transcripts V1 and V3 utilise exon 1a whilst V2 utilises 1b. Together they code for 2 different C9ORF72 protein isoforms – the short 24 kDa isoform (C9S) derived from V1 (exons 2-5) and the long 54kDa isoform derived from both V2 and V3 (exons 2-11) (Stepto et al., 2014; Haeusler et al., 2016). The G_4C_2 repeat expansion is in-between exons 1a and 1b, hence is located either within intron 1 or in the promoter depending on the transcript variant (Stepto et al., 2014; Haeusler et al., 2016). As transcripts V1 and V3 utilise exon 1a they contain the G_4C_2 sequence within intron 1 leading to the production of pre-mRNA harbouring the expansion (Stepto et al., 2014; Haeusler et al., 2016). V2 incorporates exon 1b hence the repeats are located within the predicted promotor region of V2 (Stepto et al., 2014; Haeusler et al., 2016). The discovery of the G_4C_2 expansion located in intron 1 of *C9ORF72* (figure 1.9) provided the strongest molecular link between ALS and FTD. As the expansion is so common in familial forms of both ALS and FTD and the combined ALS-FTD syndrome, in addition to its occurrence in sporadic forms of both diseases; understanding the underlying pathogenic mechanisms is essential for the future management of both diseases (Haeusler et al., 2016).

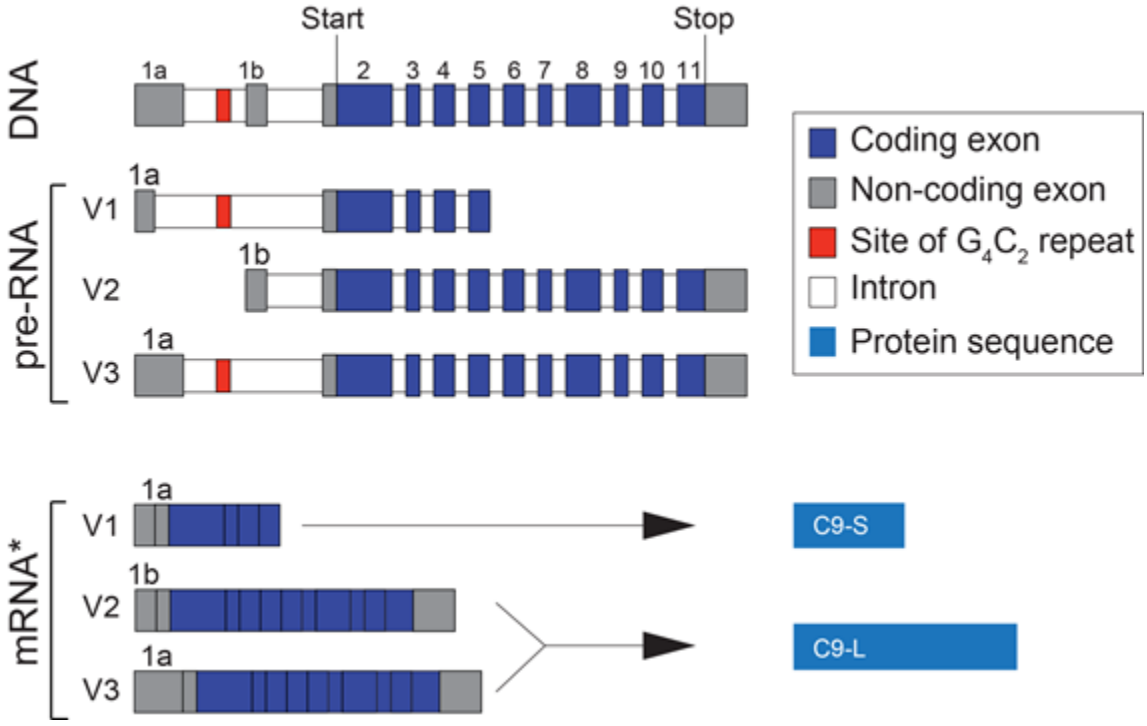

Figure 1.9 | *C9ORF72* transcript variants and protein isoforms. The *C9ORF72* gene consists of 12 exons, 10 coding exons (exons 2 – 11) and 2 non- coding exons (exons 1a and 1b). *C9ORF72* is alternatively spliced to produce 3 mRNA transcripts, V1, V2 and V3 in which the non-coding exons 1a and 1b are differentially incorporated into the pre-RNA. The G_4C_2 repeat expansion is located in-between exons 1a and 1b, hence is located either within intron 1 or in the promoter depending on the transcript variant**.** Together they code for 2 different C9ORF72 protein isoforms – the short 24 kDa isoform (C9S) derived from V1 (exons 2-5) and the long 54kDa isoform derived from both V2 and V3 (exons 2-11). * The mRNA is depicted without presence of the G_4C_2 repeats. Adapted from Stepto et al. (2014).

1.4.1 Origin of the C9ORF72 repeat expansion

An important question remains as to whether the G_4C_2 repeat expansion in *C9ORF72* has a single founder that has spread throughout the population or if the *C9ORF72* locus is prone to expansion and occurred on multiple occasions throughout human history (Pliner et al., 2014). A closer analysis of the original Finnish GWAS that identified the *C9ORF72* locus revealed all patients had a 232-kb haplotype which increased ALS incidence (Laaksovirta et al., 2010). Furthermore, the majority of non-Finnish patients also carry this 20-single nucleotide polymorphism (SNP) risk haplotype (Majounie et al., 2012b). The link between the repeat expansion in *C9ORF72* and this risk haplotype not only holds for European populations (Ratti et al., 2012, Smith et al., 2013; van der Zee et al., 2013) but also Japanese (Ishiura et al., 2012; Ogaki et al., 2012) and Chinese (Jiao et al., 2014) populations.

Genome-wide genotype data from Finland shows the G_4C_2 repeat expansion appeared in the Finnish population 100 generations ago (Majounie et al., 2012b; Pliner et al., 2014); which suggests the mutation occurred around 500 A.D., the time the Vikings invaded Europe (Pliner et al., 2014). A “Viking horde” theory has been proposed which suggests the expansion spread through Europe during the Viking conquests (Mok et al., 2012; Pliner et al., 2014). Pliner et al. (2014) describe how the geographic distribution of the *C9ORF72* repeat expansion and the conquests of the Vikings are strikingly similar. Increased globalisation will have spread the *C9ORF72* expansion to other populations, explaining the Finnish haplotype occurring in Asian expansion carriers (Pliner et al., 2014). Pliner et al. (2014) argue the geographic distribution of the *C9ORF72* expansion within Europe supports the single founder theory, as if the expansion occurred numerous times throughout human history then several areas with high numbers of *C9ORF72* disease frequency would be predicted across the globe.

Rohrer et al. (2015) note however that the high prevalence of the *C9ORF72* expansion in certain southern European populations and its occurrence in large populations in East Asia such as China do not support this “Viking horde” theory. Rohrer et al. (2015) further argue the common 20-SNP haplotype is more likely to have an out-of-Africa origin as it is observed in European, African, and Asian

populations. Rather than a single common founder the alternative hypothesis is that of a predisposing haplotype prone to mutations leading to the occurrence of multiple expansions and hence multiple origins worldwide (Rohrer et al., 2015). Several studies looking at generational changes in repeat length within families and somatic instability of the repeats support a multiple origin hypothesis for the *C9ORF72* expansion (Rohrer et al., 2015). Fratta et al. (2015) found within a UK ALS cohort a patient with a 90 repeat unit expansion in ectodermally derived cauda equine and other tissues such as fibroblasts, blood and dura mater that arise from the mesoderm and/or ectoderm. A much larger expansion of between 900 – 3500 repeats was only present in CNS tissue. The authors argue the large number of small expanded alleles similar in size across multiple non-CNS tissues makes the retraction of the large expansion improbable as this would have to involve multiple retraction events that lead to a similar final repeat length. Rather, Fratta et al. (2015) suggest this indicates the expansion has occurred multiple times on a permissive haplotype; indeed the patient with the 90 repeat unit expansion in non-CNS tissue did have the Finnish founder haplotype. Xi et al. (2015a) described a British-Canadian family with a paternal 70 repeat unit expansion allele which expanded to around ~1750 repeats in the blood of 4 of the 5 children; two of which have ALS. Similarly, Dols-Icardo et al. (2014) report a family of 3 brothers with expansions ranging between 116 – 148 in blood; with the children of two of these brothers having repeat lengths of 120 and 1401 respectively. Gijselinck et al. (2016) also observed a large increase of around ~1000 repeats from a father to his children. Hence there is strong evidence for both unfaithful inheritance (Beck et al., 2013; Dols-Icardo et al., 2014; Xi et al., 2015a) and somatic instability (van Blitterswijk et al., 2013; Fratta et al., 2015; Nordin et al., 2015) for intermediate (20-30) and small (30-150) repeat G_4C_2 expansions. Such generational increases indicate it is improbable that the repeat expansion only happened once during the history of humanity since expansion events are still occurring today.

1.4.2 Pathogenic threshold of the C9ORF72 repeat expansion

With the discovery of G_4C_2 repeats in *C9ORF72*, an important question is what the minimum repeat length required for disease is (figure 1.10). The normal repeat size in healthy controls varies but is between 2 – 10 repeats in 90% of the European

population (Renton et al., 2011). Most controls however possess 2, 5 or 8 repeats on each allele (Woollacott and Mead, 2014) although intermediate repeat lengths of 30 repeats and above are seen uncommonly in healthy populations (Rohrer et al., 2015). An expansion of 400 repeats has been reported in controls and has been attributed to reduced penetrance of the *C9ORF72* expansion and variations in age of onset (Beck et al., 2013; Mossevelde et al., 2017). Repeat expansions in patients are much larger than what is observed in healthy controls, ranging from the hundreds to the thousands (Rohrer et al., 2015; Mossevelde et al., 2017) (figure 1.10). Accurately sizing the repeat expansion is difficult due to the high GC content of the expansion (Stepto et al., 2014). Repeat prime PCR used in the discovery of the expansion cannot accurately distinguish repeat sizes larger than 30 – 50 repeats (Renton et al., 2011); although new protocols have been developed that sequence more than ~100 repeats (Cleary et al., 2016). Southern blotting is generally considered the gold standard for detecting and sizing large repeat expansions (Cleary et al., 2016); however somatic mosaicism for the repeat length in blood samples makes accurate interpretation difficult; additionally, the resolution of agarose gel electrophoresis used in southern blotting is lower than PCR and is more affected by secondary structure as the gel is non-denaturing (Cleary et al., 2016). It is recommended that the use of both PCR and southern blotting should be used for clinical diagnosis (Akimoto et al., 2014).

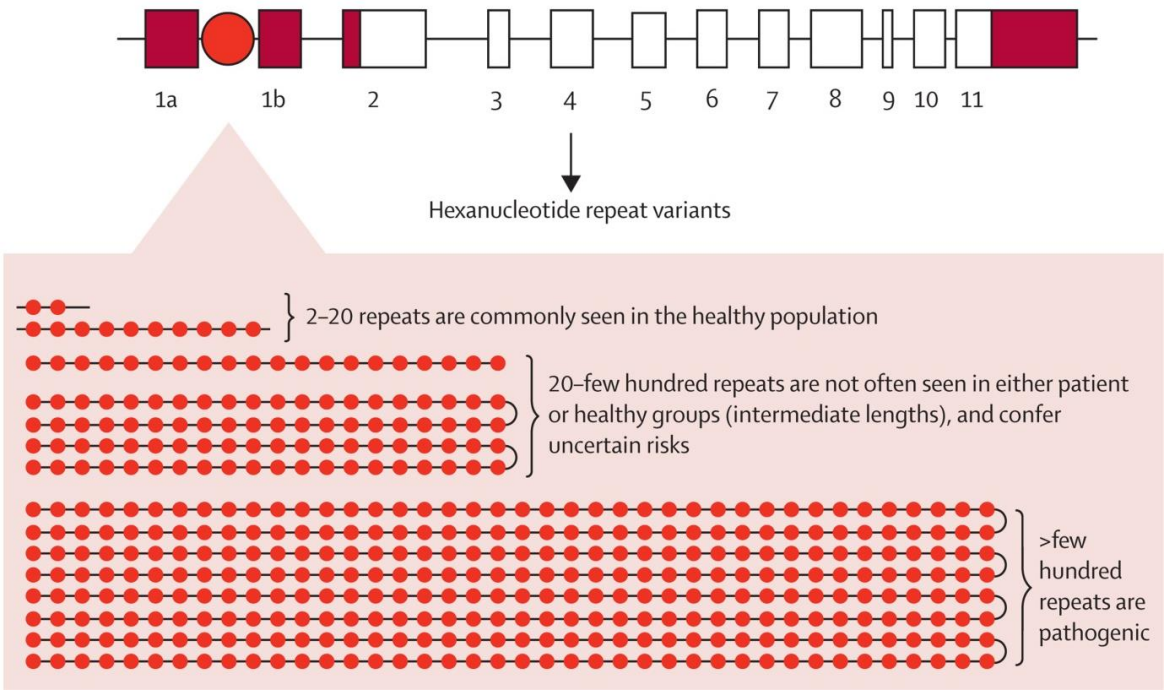

Figure 1.10 | Repeat lengths in C9orf72 associated with disease risk. The G_4C_2 repeat expansion in *C9ORF72* is in intron 1 of the gene. Repeat lengths in the healthy range are between 2-20 repeats; repeat lengths between 20 – a few hundred repeats are rare in both patients and healthy controls hence present an unknown risk for disease. Over 400 hundred repeats are pathogenic. Adapted from Rohrer et al. (2015).

There appears to be an intermediate range between 20 – a few hundred repeats that are rarely seen in healthy controls or patients that confer an unknown risk for the disease (Rohrer et al., 2015) (figure 1.10). G_4C_2 expansions ranging between 24 – 28 repeats have previously been reported in ALS patients (Milliecamps et al., 2012; Ratti et al., 2012; García-Redondo et al., 2013) and repeat lengths of 20 – 22 units have been seen in FTD patients (Gómez-Tortosa et al., 2013). Gijselinck et al. (2016) found short expansions between 47 – 78 repeat units in blood co-segregate with ALS and FTD. Some of these patients with short expansions presented with *C9ORF72* repeat expansion associated ALS/FTD pathology in the brain that was indistinguishable from the pathology seen patients with much longer expansions (Gijselinck et al., 2016). However, in non-expansion carriers there is at present no clear association between repeat lengths within the normal range and the risk of developing ALS, FTD or ALS-FTD (Mossevelde et al., 2017). Healthy individuals harbouring repeat expansions between 20 – a few hundred repeats have been

reported by a number of groups (Simón-Sánchez et al., 2012; Cooper-Knock et al., 2012; Majounie et al., 2012a; Beck et al., 2013; Smith et al., 2013; Gami et al., 2015). Gami et al. (2015) reported the case of an 84-year-old woman who had a 30-unit repeat expansion but did not develop ALS or FTD or experience any sort of significant cognitive decline. One study has observed an increased chance of developing sporadic ALS is associated with intermediate 24 – 30 repeats expansions (Chen et al., 2015). Hence whether intermediate repeat lengths are pathogenic or confer increased risk to develop clinical disease is at present unclear.

There is currently no clear evidence for a strong correlation between repeat length and disease severity and survival (Rutherford et al., 2012). No association has been found between repeat length and the clinical and pathological diagnosis of either FTD or ALS (van Blitterswijk et al., 2013; Nordin et al., 2015). Furthermore, in ALS patients repeat length did not influence bulbar or spinal presentations of the disease (Dols-Icardo et al., 2014; Chen et al., 2015). On the other hand, some have reported links between repeat length and patient survival. Suh et al. (2015) found large G_4C_2 repeat expansions correlate with reduced disease duration in FTD. Post mortem analysis of the brains of FTD and ALS patients by van Blitterswijk et al. (2013) showed a correlation between repeat length in the cerebellum and disease duration; with repeats over 1467 in length conferring reduced survival after onset in FTD. Similarly, Nordin et al. (2015) found large expansions in the parietal lobe associated with quicker disease progression in a cohort of 13 ALS patients and 5 ALS/FTD patients; this was significant for both the total cohort and the ALS patient group when analysed in isolation. However, the authors saw no correlation between length of repeat in the spinal cord and disease duration after onset (Nordin et al., 2015). Other studies have found no correlation between the size of the repeat expansion and disease duration of either ALS or FTD (Rutherford et al., 2012; Dols-Icardo et al*.,* (2014). Some indirect data suggests an inverse correlation may exist between repeat size and disease duration; hypermethylation of the CpG island located 5' of the G_4C_2 repeats correlates with both larger repeat size (Gijselinck et al., 2016) and increased duration of the disease in FTD (Russ et al., 2015). Hypermethylation of the 5' CpG island has been proposed to be protective due to epigenetic silencing which causes a reduction in mutant *C9ORF72* mRNA, and as a consequence potentially attenuating the toxic effects of the gain of function species derived from the G_4C_2

expansion (Russ et al., 2015). Indeed, reduced pathology is seen in post-mortem patient brains with CpG hypermethylation (Liu et al., 2014). However, a conflicting study by Xi et al. (2013) found increased methylation of the 5' CpG island correlated with reduced disease duration in ALS. Hence a firm association between repeat size and disease severity and duration is unclear at present.

Repeat length may be an age of onset modifier; an earlier age of onset is seen in the younger generations of *C9ORF72* families indicating genetic anticipation (Chiò et al., 2012a; Boeve et al., 2012; Gijselinck et al., (2016). Gijselinck et al. (2016) found the offspring of families showing intergenerational decreases in age of onset have a higher methylation state of the flanking CpG island compared to the parent. Hypermethylation is associated with increased repeat sizes; suggesting an increase in expansion size between generations has occurred and may explain the earlier disease onset. Indeed, in one of the families studied Gijselinck et al. (2016), elevated methylation and an increase of around 1000 repeats units was seen from parent to child. Genetic anticipation has not been universally observed in all *C9ORF72* families (Renton et al., 2011; DeJesus-Hernandez et al., 2011; Boeve et al., 2012; Chiò et al., 2012a). Indeed, Gijselinck et al. (2016) found when comparing the age of onset of 6 patients harbouring a short expansion with 51 patients with a long expansion, that 47-80 repeat units in blood was significantly associated with a later age of onset compared to 80 or more repeats. Gijselinck et al. (2016) propose this difference in age of onset maybe attributable to reduced methylation of the CpG island 5' of the repeats in both the blood and brain of patients with short repeat expansions. Conversely, Beck et al. (2013), Waite et al. (2014) and Hübers et al. (2014). all found a positive correlation between length of the expansion and age of onset, with those patients with smaller expansions having an earlier age of onset. Additionally, van Blitterswijk et al. (2013) found larger repeat lengths in the frontal cortex are significantly associated with increased age of onset. This finding was not replicated by Nordin et al. (2015) but the authors did observe a positive correlation between repeat length in parietal lobe and cerebellum and age of onset. However, Mossevelde et al. (2017) note such results must be interpreted with caution as these studies used different measures to assess repeat size which may impact the results. In addition Hübers et al. (2014) found this positive correlation is no longer significant

when an outlier with an extremely early age of onset (27 years of age) and a repeat length of 250 units was removed from the analysis.

1.4.2 Clinical presentation of patients harbouring the C9ORF72 repeat expansion

Patients harbouring the hexanucleotide G_4C_2 in *C9ORF72* typically have a rapid disease progression like the majority of ALS and ALS-FTD patients (Rohrer et al., 2015). Life expectancy does however depend on the clinical diagnosis. ALS associated *C9ORF72* repeat expansions usually has a disease duration between 3 to 96 months (Cruts et al., 2015) whereas for *C9ORF72* associated FTD disease duration can range from 1 to 22 years (Cruts et al., 2015). Atypical slowly progressive forms of *C9ORF72* associated ALS and FTD have been reported with patients living 20 years without a significant worsening of symptoms after diagnosis (Khan et al., 2012; Suhoen et al., 2014; Gómez-Tortosa et al., 2014). Age of onset is highly variable, between 30 – 70 years (range: 27 – 85 years), irrespective of how the disease initially manifests itself (Cruts et al., 2013; Cruts et al., 2015) with mean age of onset being around 58.0 (± 8 years) (Majounie et al., 2012a; van der Zee et al., 2013; Murphy et al., 2017). The expansion is typically not penetrant in those younger than 35 years of age although a 25-year-old ALS patient with the expansion has been reported (Murphy et al., 2017). Penetrance of the expansion is around 50% at around 60 years with penetrance being 99.5% by 83 years of age (Murphy et al., 2017). Penetrance by age is slightly increased in *C9ORF72* patients with pure ALS (median range: 56 – 58) compared to those presenting with pure FTD (median range: 57.0–60.0) (Murphy et al., 2017). Increased penetrance by age is also seen in patients initially presenting with a spinal-onset form (median range: 56 – 58) of ALS compared to bulbar onset (median range: 57.0–60.0) (Murphy et al., 2017). Interestingly there appears to be increased penetrance in men compared to woman (Williams et al., 2013; Murphy et al., 2017); with males who carry the repeat expansion in *C9ORF72* having an increased chance of developing disease at earlier ages (median range: 56 – 58) compared to females (median range: 86 – 68). This gender difference is seen across the entire range of age of onset (Murphy et al., 2017), although the effect on survival is relatively small (2-year difference in median survival) (Murphy et al., 2017).

C9ORF72 repeat expansion carriers typically present with bvFTD, ALS, or a combination of both (Rohrer et al., 2015). The language variants of FTD are rarely associated with the *C9ORF72* expansion (Rohrer et al., 2015). Interestingly features of psychosis (including delusions and hallucinations) are commonly observed in bvFTD patients with the expansion in *C9ORF72* compared to those without the expansion (Rohrer et al., 2015). In certain patients these symptoms are particularly prevalent during the early stages of the disease, so much so that schizophrenia (Galimberti et al., 2014), bipolar disorder (Meisler et al., 2013), obsessive-compulsive disorder (Calvo et al., 2013) or depressive pseudodementia (Bieniek et al., 2014) are misdiagnosed closer to onset. *C9ORF72* expansion carriers can also present with amnesic phenotypes such as impairments in episodic memory (Rohrer et al., 2015).

These neuropsychiatric and amnesic features combined complicate the assigning of a bvFTD diagnosis as patients are less likely to fit in with the typical diagnostic criteria seen in non-expansion carriers (Rohrer et al., 2015), potentially leading to patients incorrectly being classified as having Alzheimer's disease (Rohrer et al., 2015). *C9ORF72* repeat expansion carriers presenting with ALS demonstrate symptoms identical to classic ALS phenotypes (Rohrer et al., 2015); although cognitive and behavioural impairments are more frequently observed in ALS patients with the *C9ORF72* expansion compared to those without (Montuschi et al., 2015). Repeat expansions in *C9ORF72* have also been reported in Parkinson disease (Lesage et al., 2013), dementia with Lewy bodies (Robinson et al., 2014), multiple system atrophy (Goldman et al., 2014), Huntington disease phenocopies (Hensman et al., 2014), pure cerebellar ataxia (Corcia et al., 2016) and corticobasal syndrome (Lindquist et al., 2013). Further some FTD patients with the *C9ORF72* repeat expansion present with parkinsonian phenotypes (Rohrer et al., 2015).

1.4.3 Pathological hallmarks of C9orf72 ALS/FTD

Like nearly all ALS cases, the majority of *C9ORF72* cases have TDP-43 pathology irrespective of the clinical phenotype (Murray et al., 2011; Cooper-knock et al., 2012; Snowden et al., 2012; Mackenzie et al., 2014; Rohrer et al., 2015). Although, rare, but notable cases have been reported that lack TDP-43 aggregation or the pathology is too sparse to classify (Gijselinck et al., 2012; Proudfoot et al., 2014; Baborie et al., 2015; Vatsavayai et al., 2016). Consistent with other TDP-43 proteinopathies, TDP-43 aggregate formation in the *C9ORF72* associated ALS/FTD brain is associated with reduction in the normal diffuse nuclear TDP-43 stain (Mackenzie et al., 2014). Pathological TDP-43 is hyperphosphorylated, ubiquitinated and N-terminally truncated (Mackenzie et al., 2014). In *C9ORF72* cases the TDP-43 pathology presents as compact or granular neuronal cytoplasmic inclusions (NCI), diffuse neuronal cytoplasmic staining, dystrophic neurites (DN), (oligodendro)glial cytoplasmic inclusions (GCI) and neuronal intranuclear inclusions (NII) (Mackenzie et al., 2014). The length of the repeat expansion does not appear to influence the presentation of TDP-43 pathology (Mann et al., 2013). TDP-43 pathology is seen in a wide range of brain areas including the frontal cortex, temporal cortex, motor cortex, hippocampus, amygdala, basal ganglia regions and the thalamus (Murray et al., 2011; Hsiung et al., 2012; Mackenzie et al., 2014). Similar brain areas show TDP-43 pathology across the *C9ORF72* clinical spectrum of ALS, ALS/FTD and FTD, the severity of TDP-43 pathology in specific brain regions does differ with the different clinical phenotypes (Mackenzie et al., 2014). However, the extent of TDP-43 pathology shows a high degree of association with neurodegeneration irrespective of the anatomical region (Hsiung et al., 2012; Mackenzie et al., 2013; Mackenzie et al., 2014). The pathology of *C9ORF72 associated* ALS (C9ALS) is identical to classic ALS pathology - marked neurodegeneration, TDP-43 NCI and pre-inclusions seen mainly in the upper motor neurons, spinal motor neurons and brain stem (Mackenzie et al., 2014). *C9ORF72* carriers with FTD (C9FTD) show less neurodegeneration and TDP-43 pathology in the motor system compared to ALS cases (Mackenzie et al., 2014). In C9FTD lower motor neuron TDP-43 pathology and cell loss is attenuated compared to pure ALS and ALS/FTD cases; although TDP-43 inclusions are still present (Mackenzie et al., 2014). In the combined *C9ORF72* ALS/FTD (C9ALS/FTD) syndrome neurodegeneration and TDP-43 pathology in the frontal and

temporal cortices is more severe compared to patients with a pure ALS phenotype (Murray et al., 2011; Hsiung et al., 2012; Mackenzie et al., 2014). The TDP-43 pathology in the neocortex of *C9ORF72* brains is reminiscent of FTD-TDP type B pathology with most of TDP-43 inclusions being compact NCI present in all cortical layers (Mackenzie et al., 2014). Although there is heterogeneity with cases showing FTD-TDP type A, particularly those with pure FTD, and also a rare FTD-TDP type C pathology also being reported in the literature (Mackenzie et al., 2014)

A unique pathology to patients harbouring the *C9ORF72* repeat expansion is the presence of ubiquitin and p62 positive; but TDP-43 negative NCI, NII and DNs; particularly in the cerebellar granular layer, neocortex, hippocampus, thalamus and amygdala (Al-Sarraj et al., 2011; Troakes et al., 2012; Cooper-Knock et al., 2012; Brettschneider et al., 2012; Hsiung et al., 2012; Mahoney et al., 2012; Mori et al., 2013a; Ash et al., 2013; Zu et al., 2013; Mackenzie et al., 2014; Schludi et al., 2015). These TDP-43 negative inclusions are the dipeptide repeat proteins translated from the G_4C_2 repeat expansion via an unconventional mechanism known as repeat-associated non-ATG initiated (RAN) translation (Mori et al., 2013a; Ash et al., 2013; Mackenzie et al., 2014; Schludi et al., 2015). The repeats are translated from both the sense G_4C_2 strand (Mori et al., 2013a; Ash et al., 2013) and the antisense C_4G_2 strand (Gendron et al., 2013; Mori et al., 2013b; Zu et al., 2013).

Translation occurs in all reading frames, the 3 sense frames and the 3 anti-sense frames - generating 5 different poly-dipeptide species each of which consists of two amino acids in a repetitive sequence (Mackenzie et al., 2014). The 5 DPRs are – poly glycine-proline (poly-GP), poly glycine-alanine (poly-GA), poly-glycine-arginine (poly-GR), poly proline-alanine (poly-PA) and poly proline-arginine (poly-PR) (Mori et al., 2013a; Ash et al., 2013; Gendron et al., 2013; Mori et al., 2013b; Mackenzie et al., 2014; Schludi et al., 2015). Poly-GP is transcribed from both the sense and anti-sense strands (Gendron et al., 2013; Mori et al., 2013b; Zu et al., 2013), poly-GA and poly-GR are produced from the sense G_4C_2 mRNA (Mori et al., 2013a; Ash et al., 2013) whereas poly-PA and poly-PR result from translation of the anti-sense C_4G_2 mRNA. (Gendron et al., 2013; Mori et al., 2013b). Additionally, presumed ATG start codons exist within the poly-PR and poly-GP anti-sense reading frames (Zu et al., 2013). RAN proteins from 5 of the 6 reading frames are predicted to have a unique C-terminal region assuming translation terminates at the first STOP codon

after the repeat sequence; this is not the case for the anti-sense poly-GP frame as a STOP codon is present directly after the repeats (Zu et al., 2013). Indeed Zu et al., (2013) developed unique C-terminal polyclonal antibodies for the RAN peptides in addition to generating antibodies to the repeat motifs; the authors were able to show sense and antisense RAN proteins accumulate in the *C9ORF72* post mortem brain using both the repeat motif and C-terminal DPR antibodies (Zu et al., 2013).

Investigation of DPR pathology in human tissue has shown sense derived DPRs are more abundant than their anti-sense counterparts; with poly-GA being the most abundant DPR followed by poly-GP and then poly-GR (Mori et al., 2013a; Mann et al., 2013; Schludi et al., 2015); with poly-PA and poly-PR pathology being rare (Gendron et al., 2013; Mori et al., 2013b; Zu et al 2013; Schludi et al., 2015). Although this appears not the be the case for poly-GP, by performing double immunofluorescence using a pure repeat motif antibody which detects both the sense and anti-sense protein and an antibody that detects the C-terminal of the sense poly-GP peptide the authors demonstrated that most of the poly-GP inclusions are produced from the anti-sense C_4G_2 mRNA strand. Double-labelling revealed two different inclusion types, sense inclusions labelled by both antibodies and anti-sense inclusions labelled only with the pure repeat motif antibody. Approximately 18% of inclusions were sense (detected by both pure repeat motif and C-terminal antibodies) whereas 82% of inclusions were antisense (detected only by the pure repeat motif antibody) (Zu et al., 2013). However more recently Jiang et al. (2016) using anti-sense oligonucleotides (ASOs) against human *C9ORF72* in a transgenic mouse model reduced poly-GP to almost undetectable levels suggesting at least in this murine model poly-GP is primarily translated from the sense strand.

DPR inclusions are primarily found in neuronal populations (figure 1.11) and are virtually always p62 positive with double labelling revealing DPR inclusions can consist of co-aggregates of more than one DPR species (Mori et al., 2013a). Cortical DPR NCIs can appear as granular dot-like aggregates in small non-pyramidal neurons to larger star shaped aggregates in pyramidal neurons (Mackenzie et al., 2014). Diffuse cytoplasmic DPR staining has also been reported; suggestive of pre-inclusions (Mackenzie et al., 2014). The majority of DPR inclusions are NCIs however NIIs and DNs, although less abundant, are observed (Mackenzie et al., 2014; Schludi et al., 2015). The majority of NIIs are p62-positive and aggregate in a

para-nucleolar compartment but do not aggregate within the nucleolus itself (Schludi et al., 2015). DPRs have also been reported in glia (Schludi et al., 2015). DPR NIIs have been observed in the ependymal cells of the spinal cord central canal and ependymal and subependymal cells that line the lateral ventricle in C9ALS cases using antibodies for poly-GA, poly-GP and poly-GR (Schludi et al., 2015). Very few of the DPR glial inclusions were cytoplasmic unlike what is seen for neuronal DPR pathology (Schludi et al., 2015). Furthermore, unlike DPR NIIs found in neurons, these glial inclusions were not para-nucleolar (Schludi et al., 2015). There is no clear association between DPR and TDP-43 pathology; indeed, DPR and TDP-43 inclusions are seldom seen in the same cell and their pathologies appear distinct (Mackenzie et al., 2014) However poly-GA has been found to surround TDP-43 aggregates (Mori et al., 2013a; Mackenzie et al 2013) suggesting that DPR aggregation may precede TDP-43 pathology.

Schludi et al., (2015) performed an extensive analysis of DPR inclusions in 36 CNS regions to analyse the extent of DPR NCI, NII and DN pathology. As was previously reported, Schludi et al., (2015) observed NCI, NII and DN pathology was most severe for poly-GA, followed by poly-GP and poly-GR. The most extensive DPR pathology was seen in the neocortex, followed by the hippocampus and cerebellum, with a high DPR burden being seen also in the amygdala and hippocampus. DPR pathology was sparse in the basal ganglia, brain stem and spinal cord. The majority of DPR inclusions are NCI; although a large amount of DNs are seen in the molecular layer of the cerebellum for poly-GA and poly-GP. Poly-GR shows less DNs as compared to the other sense DPRs; indeed, the authors note there is almost a complete lack of poly-GR DNs in the CNS. Poly-GR NIIs were also less frequent than poly-GA and poly-GP NIIs; although were more prevalent in the thalamus compared to poly-GR NCIs. No pathological analysis was performed for anti-sense DPRs as the number of inclusions were too low for accurate analysis; although the authors note the pathological distribution of poly-PR resembled what was reported for the sense DPRs (Schludi et al., 2015). Although Poly-PR inclusions were rare they were reported with highest frequency in the hippocampus (Schludi et al., 2015). Poly-PR was not found in the spinal cord contradicting previous reports of poly-PR inclusions in spinal motor neurons (Wen et al., 2014). Intriguingly poly-PR, but not poly-GR inclusions show a different distribution in FTD cases compared to ALS

(Schludi et al., 2015). Schludi et al. (2015) counted inclusions in three neocortical regions (medial frontal gyrus, motor cortex and the occipital cortex), two hippocampal regions (granular cell layer of the dentate gyrus, pyramidal cell layer of cornu ammonis regions 3 and 4) and the granular and molecular layers of the cerebellum. Poly-GR pathology was similar in C9FTD, C9ALS and C9ALS/FTD patient subgroups, however poly-PR was observed significantly more frequently in the CA3/4 region of C9FTD cases compared to C9ALS and C9ALS/FTD cases. As noted previously there is an apparent lack of DPR pathology in spinal motor neurons leading some to question the pathogenic contribution of the DPRs in C9ALS and C9ALS/FTD cases (Gomez-Deza et al., 2015; Mackenzie et al., 2015). Indeed, the pathological distribution of DPR proteins that can be visualised with current antibodies does not correlate with areas of the brain that undergo the most significant loss of neurons in C9ALS/FTD (Mackenzie et al., 2015). A summary of the distribution of sense DPR pathology in the brain is presented in figure 1.12.

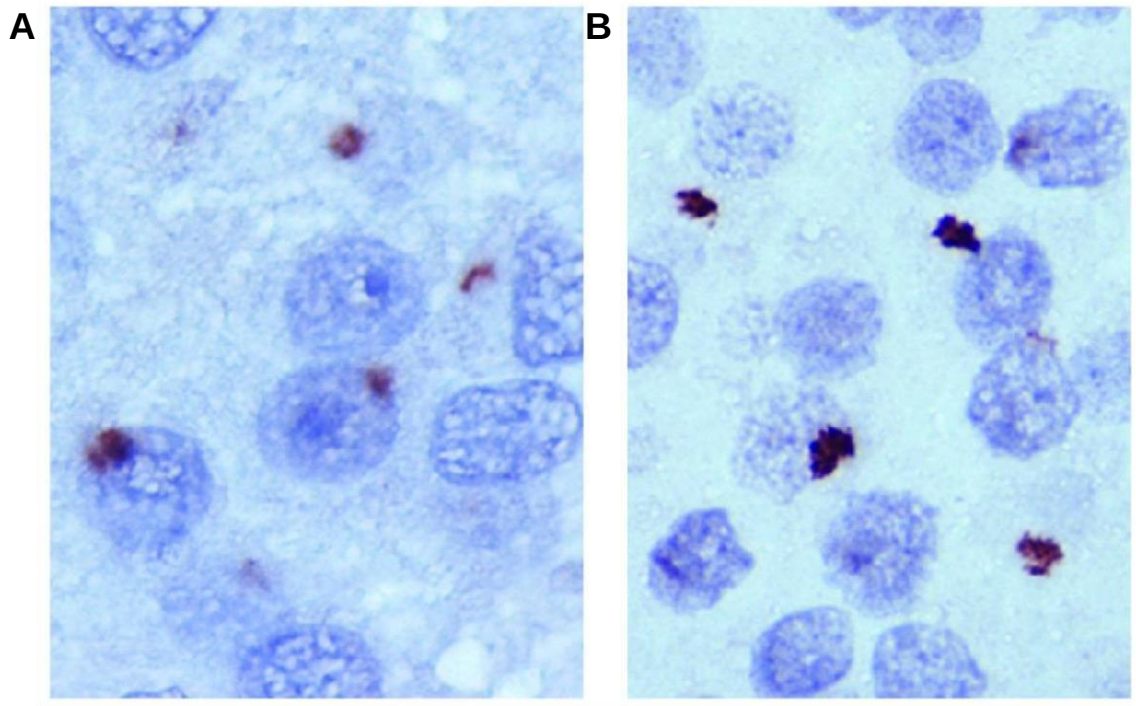

Figure 1.11 | DPR pathology in *C9orf72* ALS/FTD. (A) Poly-GR pathology in the dentate nucleus of a C9ALS/FTD patient. **(B)** Poly-GA pathology in the dentate nucleus of a C9ALS/FTD patient. Adapted from Friebaum & Taylor, 2017.

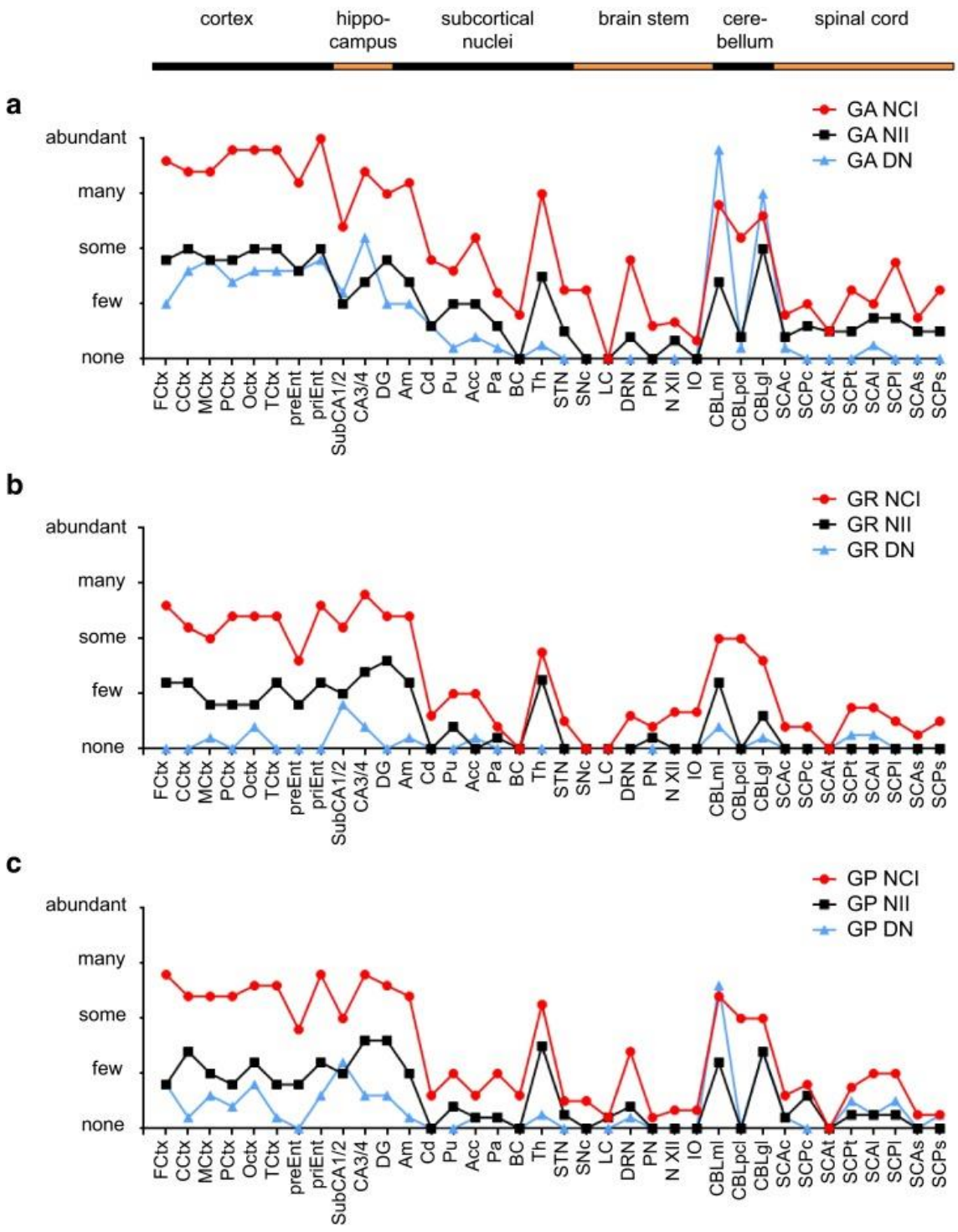

Figure 1.12 | Distribution of sense DPR pathology in different regions of the brain. (a, b & c) Poly-GA is the most abundant DPR species seen in the human C9 brain followed by poly-GP and poly-GR. Poly-GA, poly-GP and poly-GR neuronal cytoplasmic inclusions (NCI), neuronal intranuclear inclusions (NII) and dystrophic neurites (DN) are primarily seen in cortical areas, the hippocampus, amygdaloid nuclei, thalamus and cerebellum. DPR inclusions are rare in the spinal cord. From Schludi et al. (2015).

The second unique feature of the *C9ORF72* brain is the presence of G_4C_2 RNA foci in brain tissue (De-Jesus Hernandez et al., 2011; Gendron et al., 2013; Lagier-Tourenne et al., 2013; Mizielinska et al., 2013; Mackenzie et al., 2014; De-Jesus Hernandez et al., 2017) (figure 1.13). The G_4C_2 DNA is transcribed from both the sense and anti-sense strand; hence both sense and anti-sense G_4C_2 RNA foci are detected in patient brains in addition to lymphoblasts, fibroblasts and patient derived induced pluripotent stem cells (IPSCs) (Almedia et al., 2013; Donnelly et al., 2013; Lagier-Tourenne et al., 2013). RNA foci are most abundant in neurons but are observed to a lesser extent in astrocytes, oligodendrocytes and microglia (Lagier-Tourenne et al., 2013; Mizielinska et al., 2013). The vast majority of RNA foci are nuclear although rare cytoplasmic RNA foci have been reported (Mackenzie et al., 2014). RNA foci have been observed in several brain regions including the frontal cortex, temporal lobe, hippocampus, cerebellum, thalamus and lower motor neurons (De-Jesus Hernandez et al., 2011; Lagier-Tourenne et al., 2013; Mizielinska et al., 2013; Lee et al., 2013; Cooper-Knock et al., 2015; Vatsavayai et al., 2016). Sense RNA foci are more abundant than anti-sense foci within regions of the brain (Mizielinska et al., 2013; Cooper-Knock et al., 2015; De-Jesus Hernandez et al., 2017); although there are notable exceptions such as motor neurons (Cooper-Knock et al., 2015).

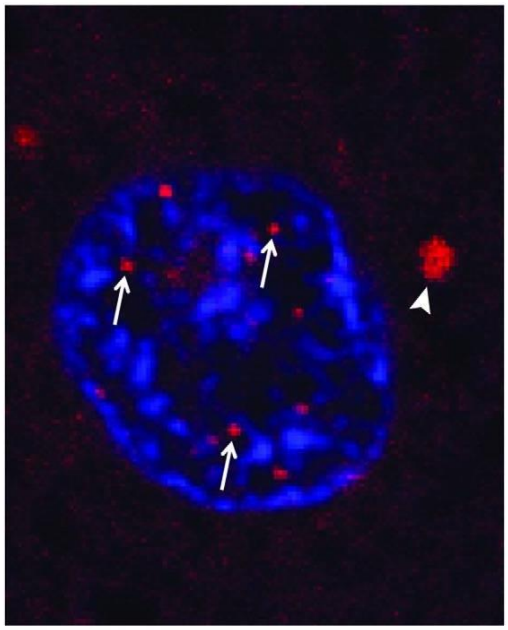

Figure 1.13 | RNA foci in *C9ORF72* expansion carriers. Neuronal sense G_4C_2 RNA foci in the frontal cortex, labelled with a Cy3-labelled (G4C2)4 oligonucleotide probe (red). Arrows denote nuclear foci; arrow heads denote cytoplasmic foci. (Adapted from Freidbaum and Taylor, 2017).

A recent study has conducted an extensive clinico-pathological study (n=63 brains) of sense and anti-sense RNA foci in frontal cortex and cerebellum (De-Jesus Hernandez et al., 2017). In frontal cortex sense RNA foci were present in 26% of cells, with 32% of sense foci bearing cells being neurons; as has previously been reported anti-sense foci were much less abundant – present in only 12% of frontal cortex cells with 16% of these being neurons. Within the cerebellum 23% of granule cells harboured sense RNA foci; sense foci were much more abundant in Purkinje cells; being found in 70% of these neurons. Anti-sense RNA foci were only found in 1% of granule cells however again were markedly present in Purkinje cells – 74% of Purkinje cells were found to contain anti-sense foci. Usually 1, 2 or 3 foci were found per cell – although there were exceptions with some cells containing a larger number of foci (De-Jesus Hernandez et al., 2017). Interestingly when anti-sense foci were found, in both the frontal cortex and cerebellar Purkinje cells, they were more common than sense foci; this was not the case for cerebellar granular cells (De-Jesus Hernandez et al., 2017). In the hippocampal neurons no difference was reported been sense and anti-sense foci burden (Cooper-Knock et al., 2015). Interestingly Cooper-Knock et al. (2015) reported anti-sense foci are more prevalent in motor neurons than their sense counterparts; further the authors also found anti-sense foci correlated with TDP-43 mislocalisation in motor neurons. This data suggests the different foci patterns seen in distinct neuronal populations may potentially mean there are differences in the contribution of sense and anti-sense foci to pathogenesis (Cooper-Knock et al., (2015). Bacterial artificial chromosome (BAC) transgenic mice show an accumulation of anti-sense foci in vulnerable regions of the CNS including layers II/III and V of the cerebral cortex, DG and CA regions of the hippocampus and cerebellar Purkinje neurons; whereas sense foci were seen widespread throughout the mouse brain (Liu et al., 2016c). In contrast to Cooper-Knock et al. (2015), only sense foci were found in motor neurons leading the authors to suggest potentially that motor neurons are more susceptible to sense foci toxicity (Liu et al., 2016c). Additionally, RNA foci and DPRs are seldom seen in the same cell (Gendron et al., 2013) although RNA foci are reported in p62 and TDP-43 bearing neurons respectively (Mizielinska et al., 2013).

1.5 C9ORF72 protein function

The C9ORF72 transcripts code for 2 different C9ORF72 protein isoforms – a short 24 kDa isoform (C9S) derived from transcript V1 and a longer 54kDa isoform which is derived from transcripts V2 and V3 (Stepto et al., 2014; Haeusler et al., 2016). Early bioinformatics analysis revealed C9ORF72 is a distant evolutionary relative of proteins of the Differentially Expressed in Normal and Neoplasia (DENN) family, GDP/GTP exchange factors (GEFs) that activate the Rab family GTPases (Zhang et al., 2012; Levine et al., 2013). The protein is highly conserved throughout evolution suggesting a common eukaryotic ancestor– with some notable exceptions including insects such as *Drosophila Melanogaster* (Stepto et al., 2014). C9ORF72 bares structural homology to DENN proteins which consist of an N-terminal longin domain, a DENN domain and a C-terminal alpha domain (figure 1.14) (Zhang et al., 2012; Levine et al., 2013).

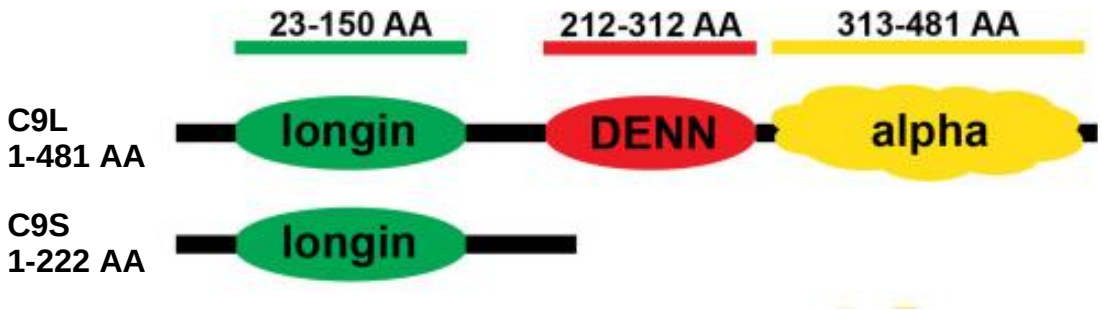

Figure 1.14 | The structure of C9ORF72 protein isoforms. The C9ORF72 protein isoforms bare structural homology to DENN proteins. The long isoform, C9L, contains an N-terminal longin domain, a DENN domain and a C-terminal alpha domain. The short isoform (C9S) contains only the N-terminal longin domain. Adapted from Xiao et al. (2016).

C9ORF72 does not contain similarity to other DENN domains at the primary amino acid level (Amick and Ferguson, 2017), its putative DENN domain was identified by the structural similarity it shares with FLCN interacting proteins 1 and 2 (FNIP1/2) and their interacting partner FLCN; both FLCN and FNIP1/2, despite not harbouring amino acid similarity with other DENN proteins, have been shown to contain DENN domains (Zhang et al., 2012; Levine et al., 2013; Amick and Ferguson 2017). DENN proteins function as GEFs for Rab GTPases, a family of small GTPases which are

master regulators cellular of membrane trafficking (Stenmark, 2009; Hutaglung and Novick, 2011). Rab GTPases exist in two conformation states – an “on” GTP bound state and an “off” GDP bound state (Stenmark, 2009). The transition between these states is controlled by Rab regulatory proteins – GEFs, GDP dissociation inhibitors (GDIs) and GTPase activating proteins (GAPs) (Stenmark, 2009; Hutaglung and Novick, 2013; Amick and Ferguson, 2017). GEFs bind to nucleotide-free GTPases causing a conformational switch leading to bound GDP being swapped for GTP (Amick and Ferguson, 2017). GEFs cause Rab GTPase proteins to be enriched at membranes where they recruit numerous molecules to regulate membrane trafficking (Amick and Ferguson, 2017). The predicted DENN domain of C9ORF72 suggested the protein may function as a GEF and hence the recruitment Rabs and associated molecules to membranes. DENN proteins mainly activate endocytotic Rabs (Cherfils and Zeghouf, 2013) indicating C9ORF72 may play a role in endosomal trafficking events, including the regulation and degradation of proteins by autophagy (Farg et al., 2014).

1.5.1 C9ORF72 and endosomal trafficking

Indeed, initial studies by Farg et al. (2014) demonstrated a role for C9ORF72 in endosomal trafficking. Using a construct encoding a GFP-tagged human C9ORF72, Farg et al. (2014) observed C9ORF72 formed vesicles in the nucleus and cytoplasm that strongly co-localised with Rab1, Rab7, Rab5 and Rab11 in both neuronal cell lines and cultured cortical neurons from mice. Furthermore Rab5, Rab7 and Rab11 were all found to co-localise with C9ORF72 in human spinal cord motor neurons. These Rabs are all associated with autophagy and endosomal transport (Farg et al., 2014). Using a siRNA to C9ORF72 Farg et al. (2014) found cells in which human C9ORF72 had been silenced the transport of the endocytosis was impaired. Aoki et al. (2017) who found a novel interaction between C9ORF72 and RAB7L1 and the loss of the C9ORF72 interaction with RAB7L1 resulted in impaired trans-Golgi network trafficking and defective multivesicular endosome formation.

C9ORF72 also regulates autophagy in neuronal cell lines (Farg et al., 2014); C9ORF72 punctate were found to co-localise with the autophagosome marker microtubule-associated protein 1 light chain 3 (LC3) and also lysosomes. Further autophagosome formation was also impaired using a C9ORF72 siRNA.

Furthermore, a dysfunctional trans-Golgi network, such as is observed when C9ORF72 is knocked down (Farg et al., 2014; Aoki et al., 2017) is known to cause disturbances in autophagy and lysosomal protein degradation pathways (Eskelinen and Saftig, 2009). This data suggest C9ORF72 participates in the macroautophagy pathway; an autophagic process in which substrates, such as misfolded proteins and damaged organelles are sequestered within a cytoplasmic double-membrane vesicle labelled autophagosomes, these autophagosomes then fuse with the lysosome where most of the degradation occurs (Feng et al., 2014). The degradation process can be divided into four steps; (i) translocation and initiation, (ii) elongation and cargo recruitment, (iii) completion and (iv) lysosome fusion and degradation (Rubinsztein et al., 2015).

1.5.2 Functional role for C9ORF72 in macroautophagy and lysosomal biology

The role of C9ORF72 in macroautophagy (herein referred to as autophagy) has been clarified by a series of studies by Webster et al. (2016a) and Sellier et al. (2016). Webster et al. (2016a) demonstrated C9ORF72 participates in the initiation step of the macroautophagy pathway, by acting as a Rab1a effector. Webster et al. (2016a) using two siRNAs that targeted both isoforms (C9S and C9L) found a reduction in autophagosome formation. Webster et al. (2016a) go on to show that overexpression of both C9S and C9L isoforms induced autophagy, and this induction was dependent on its interaction with the ULK1 autophagy initiation complex. This complex consists of Unc-51-like kinase 1 (ULK1), FAK family kinase-interacting protein of 200 kDa (FIP200), autophagy-related 13 (ATG13) and ATG101 (Webster et al., 2016b). Using a FIP200 siRNA to disrupt the ULK1 complex it was found overexpression of C9ORF72 no longer promoted the induction of autophagy. Both co-immunoprecipitation and proximity ligation assays demonstrated the C9S and C9L isoforms directly interact with ULK1, FIP200 and ATG13. C9ORF72 did not seem interact with the ULK1 complex at the level of mTOR; rather C9ORF72 was shown to regulate the translocation of the ULK1 complex to the phagophore by acting as a Rab1a effector.

Further insight into how C9ORF72 regulates autophagy has come from a complimentary study published in conjunction with Webster et al. (2016a) by Sellier

et al. (2016). Sellier et al. (2016) demonstrated that C9ORF72 forms a complex with WDR41 and SMCR8; this C9ORF72, SMCR8, and WDR41 complex acts as a GDP/GTP exchange factor for RAB8a and RAB39b; with the GEF activity of the C9ORF72 protein complex with RAB39b being crucial for the modulatory role C9ORF72 plays in autophagy in neurons. This protein complex was found by Sellier et al. (2016) to only be active in the presence of SMCR8. Indeed, SMCR8 appears vital for the stability of C9ORF72 and vice versa; levels of both proteins are decreased markedly when either is knocked down (Amick et al., 2016; Ugolino et al., 2016). Yang et al. (2016a) demonstrated C9ORF72 interacts with SMCR8 in a DENN-domain dependent manner; as deletion of the SMCR8 DENN domain completely removed its interaction with C9ORF72. Further investigation of this emerging C9ORF72, SMCR8, and WDR41 pathway by Sellier et al. (2016) revealed TANK-Binding Kinase 1 (TBK1) phosphorylates SMCR8, and this post-translational modification is significant for the regulation of autophagy. Sullivan et al. (2016) demonstrated the C9ORF72/SMCR8/WDR41 forms a ternary complex, with FIP200 identified by Webster et al. (2016a) as a C9ORF72 binding partner; this suggests the formation of the C9ORF72/SMCR8/WDR41 complex enables their interaction with the FIP200/Ulk1/ATG13/ATG101 autophagy initiation complex (Sullivan et al., 2016).

Related to autophagy; C9ORF72 and SMCR8 also appear to play a role in lysosomal biology (figure 1.13). Both proteins localise to the lysosome during cellular starvation (Amick et al., 2016) – although are not present at the lysosome under basal conditions (Webster et al., 2016a; Yang et al., 2016a). Following knockout of SMCR8 and C9ORF72 the lysosomal deficits have been reported - including enlarged lysosomes, increased swollen perinuclear lysosomes, and build-up of lysosomes and lysosomal enzymes (O'Rourke et al., 2016; Amick et al., 2016; Sellier et al., 2016; Sullivan et al., 2016; Ugolino et al., 2016; Nassif et al., 2017). Nassif et al. (2017) note this suggests C9ORF72 might play a role in the Coordinated Lysosomal Expression and Regulation (CLEAR) network involved in transcription of genes involved in both lysosomal biogenesis and macroautophagy. The transcription factor EB (TFEB) plays a key role in CLEAR, and phosphorylation by mTORC1 is an important inhibitor of TFEB nuclear translocation (Nassif et al., 2017). C9ORF72 knockout cells show mTORC1 inhibition under starvation conditions or amino acid deprivation (Amick et al., 2016; Ugolino et al., 2016). In C9ORF72 knockout cells,

mTORC1 inhibition leads to increased translocation of TFEB to the nucleus and an enhancement of autophagy flux (Ugolino et al., 2016).

In summary a pathway is emerging in which C9ORF72 regulates the initiation of autophagy by playing a central role in the RAB cascade pathway (figure 1.15) (Webster et al., 2016b). C9ORF72 functions as a RAB1a effector; interacting with the ULK1 complex and active GTP-RAB1a thereby controlling the translocation of the ULK1 initiation complex to the site of phagophore formation (Webster et al., 2016a; Nassif et al., 2017). Once at the site of phagophore formation C9ORF72 forms a complex with SMCR8 and WDR41 where it acts as a GEF for RAB8a and RAB36b (Sellier et al., 2016). Active RAB8a and RAB36b are potentially involved in the delivery of additional membrane to the site of phagophore formation and further promoting ULK1 translocation (Webster et al., 2016b; Nassif et al., 2017). RAB8a and RAB36b are known to interact with the autophagy receptors p62 and optineurin (Sellier et al., 2016; Webster et al., 2016b); their activation via C9ORF72 may promote nucleation and the recruitment of autophagic substrates such as misfolded proteins to the phagophore (Webster et al., 2016b). TBK1 phosphorylation of SMCR8 enhances the GEF activity of C9ORF72 encouraging its interaction with RAB36b (Sellier et al., 2016). Following C9ORF72 mediated RAB dependent trafficking of the ULK1 complex, the inhibitory phosphorylation of ULK1 via mTORC1 is lost allowing the ULK1 complex to interact with the FIP200 complex and initiate autophagy. Hence C9ORF72 plays a central role in a Rab dependent trafficking pathway involved in the initiation of autophagy. C9ORF72 may also influence autophagic influx via its modulation of mTORC1 signalling and TFEB when localised to the lysosome.

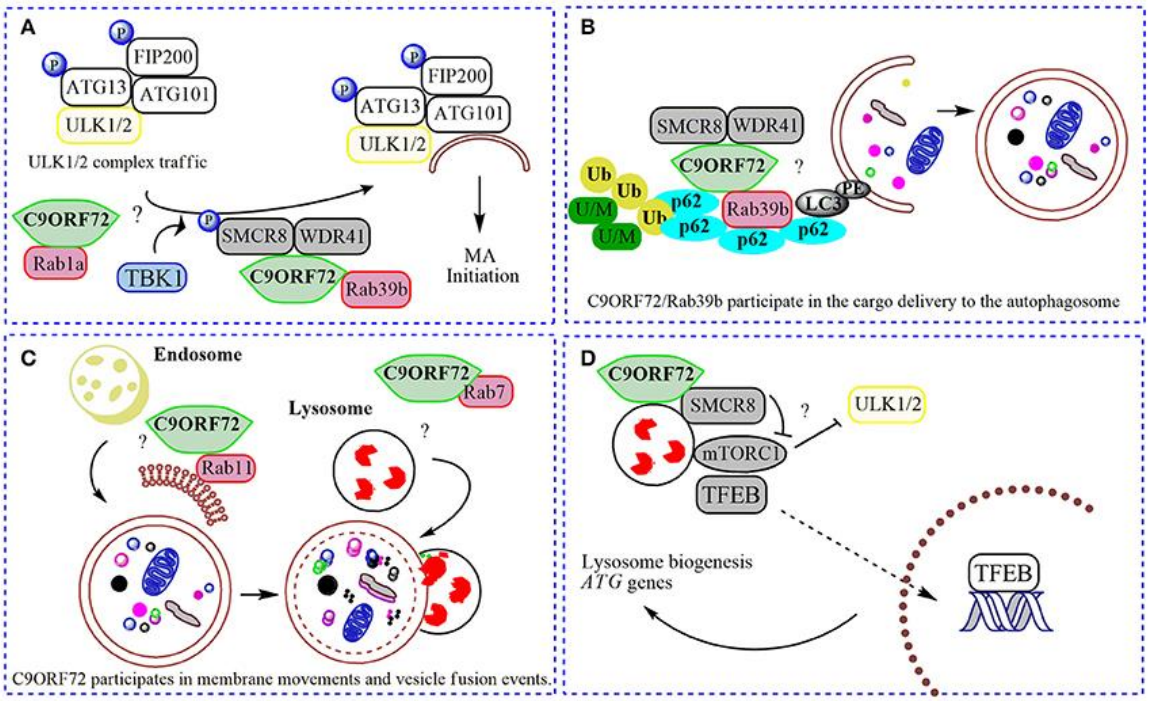

Figure 1.15 | Role of C9ORF72 in macroautophagy and lysosomal biology. **(A)** C9ORF72 functions as a Rab1a effector; interacting with the ULK1 complex and active GTP-Rab1a controlling ULK1 initiation complex translocation to the site of phagophore formation. C9ORF72 then forms a complex with SMCR8 and WDR41 where it acts as a GEF for RAB8a and RAB36b (Sellier et al., 2016). The phosphorylation of TBK1 via SMCR8 enhances the GEF activity of C9ORF72. Following C9ORF72 mediated delivery of the ULK1 complex inhibitory phosphorylation of ULK1 by mTORC1 is lost allowing the ULK1 complex to interact with the FIP200 complex to initiate autophagy. **(B)** RAB8a and RAB36b are known to interact with autophagy receptors such as p62; C9ORF72 may promote the recruitment of autophagic substrates to the phagophore (Sellier et al., 2016). **(C)** C9ORF72 also interacts with RAB11 (Farg et al., 2014) RAB7 and RAB7L1 (Farg et al., 2014; Aoki et al., 2017) hence may also be involved in membrane trafficking to the phagophore and fusion of the autophagosome and lysosome. **(D)** Finally, both C9ORF72 and SMCR8 may modulate mTORC1 nutrient sensing promoting TFEB nuclear translocation increasing lysosomal biogenesis and removing ULK1 mTORC1 inhibition of ULK1 thereby initiating autophagy. From Nassif et al. (2017).

1.5.3 Other functions of C9ORF72

A role for C9ORF72 in actin dynamics has come from studies of the mouse homologue of C9ORF72. The mouse C9ORF72 homologue (Ms-C9orf72) is known as 3111004O21Rik (Atkinson et al., 2015). The mouse homologue has 3 transcript variants; V1 (exons 2-11) encoding isoform 1 (Ms-C9orf72-1); V2 (exons 3-11) encoding isoform 2 (Ms-C9orf72-2) and V3 (exons 2-10) encoding isoform 3 (Ms-C9orf72-3) (Xiao et al., 2016). Ms-C9orf72-1 and Ms-C9orf72-3 have a highly similar domain structure to C9L, containing the longin, DENN and alpha domains; although Ms-C9orf72-3 misses 60 amino acids from the C-terminal of the peptide (Xiao et al., 2016). Ms-C9orf72-2 contains both the DENN and alpha domains but not the longin domain, hence differing from human C9S which contains only the longin domain (Xiao et al, 2016).

Atkinson et al. (2015) demonstrated in cultured cortical neurons the mouse homologue of C9ORF72 was found extensively within the microtubule cytoskeleton including the axon, soma and dendritic arbor and actin-rich structures including growth cones and filopodia (Atkinson et al., 2015). A role for C9ORF72 mechanistically in actin dynamics was provided by Sivadasan et al. (2016). Sivadasan et al. (2016) either overexpressed or knocked down Ms-C9orf72 in primary mouse embryonic motor neurons. Axons were found to be longer in motor neurons overexpressing human C9ORF72; whilst knockdown of Ms-C9orf72 led to a shortening of the axon. This knockdown phenotype could be rescued by overexpressing human C9ORF72. Further mass spectrometry-based proteomics revealed C9ORF72 interacted with proteins that regulate actin dynamics such as cofilin, Arp2/3 and coronin (Flynn et al., 2012).

Ms-C9orf72 knockout mice show autoimmunity and inflammation phenotypes (O'Rourke et al., 2016; Atanasio et al., 2016; Burberry et al., 2016; Sudria-Lopezet al., 2016; Ugolino et al., 2016; Sullivan et al., 2016). Ms-C9orf72 null mice develop progressive splenomegaly and lymphadenopathy in addition to defective lysosomal trafficking and altered immune responses in macrophages and microglia (O'Rourke et al., 2016). Atanasio et al. (2016) observed myeloid expansion, T cell activation, increased plasma cells and enhanced autoantibodies in null mice. Burberry et al.

(2016) observed similar phenotypes including increased expression of inflammatory cytokines and severe autoimmunity. Atanasio et al. (2016) suggest based on these findings that C9ORF72 regulates immune homeostasis.

C9ORF72 appears to also play a role in stress granule (SG) formation. Maharjan et al. (2017) found C9ORF72 is predominantly nuclear and localised to P-Bodies in neuronal and non-neuronal cell lines. C9ORF72 was found to colocalise with SGs in response to stressors such as ER stress and heat shock. Furthermore, C9ORF72 was shown to be required for SG formation via knockdown of both C9ORF72 isoforms using the CRISPR/Cas9 system. C9ORF72 knockdown inhibited the formation of SGs and reduced the phosphorylation of EIF2alpha which is required for stress granule assembly. Likewise, reductions in the levels of proteins involved in SG assembly were also observed; including downregulation of TIA-1, HuR and G3BP1. This reduction was also seen at the transcript level for TIA-1 and HuR suggesting C9ORF72 regulates the transcription of SG proteins.

1.5.4 C9ORF72 protein isoforms

A contentious issue in the field related to the function of the C9ORF72 protein is distinguishing between the two isoforms, C9L and C9S, and whether they have differing or complimentary roles in the processes just described. This is not helped by the lack of commercially available quality antibodies for C9ORF72 that can accurately detect the protein and distinguish between the two isoforms (Webster et al., 2016a). Structurally the two isoforms are different; the short isoform does not contain the DENN or alpha domains, only the longin domain (Xiao et al., 2016). The lack of the DENN domain in the short isoform would suggest it is the long isoform which regulates autophagy, as SMCR8 and C9ORF72 have been shown to interact by their DENN domains. Yang et al. (2016a) found in their gel filtration assay that C9L, but not C9S, comigrated with SMCR8, WDR41, and ATG101. Interestingly Sullivan et al. (2016) observed only C9L, and not C9S binds SMCR8; Xiao et al. (2016) also found in preliminary mass spectrometry and co-immunoprecipitation that only C9L bound to SMCR8 and WDR41. Taken together this data does suggest C9L may play a more important role in regulating autophagy than C9S.

Xiao et al. (2015a) are currently the only group to have developed and published antibodies specific for either of the isoforms. The isoform specific antibodies were generated by taking advantage of a unique lysine residue in the C-terminus of C9S encoded by the partial retention of intron 5 (Xiao et al., 2015b). Antibodies were generated for the C-terminal of C9L and the last 12 amino acids of C9S which includes a unique lysine residue. Both isoforms show a similar staining pattern in HEK293T cells and N2a cells, where they are found to localise diffusely in the cytoplasm and the nuclear envelope (Xiao et al., 2016). In the post-mortem human brain however labelling of the isoforms is strikingly different in neurons (Xiao et al., 2015a). In control tissue C9L was mainly diffuse in the cytoplasm of spinal motor neurons and cerebellar Purkinje neurons, however in these cerebellar Purkinje neurons speckle-like structures in the cytoplasm and dendrites were also observed. The identity of these speckles was not discovered, although they did not co-localise with endosomal, lysosomal, autophagosome or polysomal markers (Xiao et al., 2015a). C9S on the other hand had a completely different localisation, C9S was observed to localise to the nuclear membrane of Purkinje and spinal motor neurons. The localisation of C9S to the nuclear membrane suggests a potential role in nucleocytoplasmic transport and indeed C9S was found to co-localise with Lamin B, Importin-β1 and RAN-GTPase in lumbar spinal cord tissue. Hence C9ORF72 may have role in shuttling proteins across the nuclear membrane (Xiao et al., 2016).

Maharjan et al. (2017) found the interaction between SGs and C9ORF72 was specific for the long isoform of the protein. Cells were transfected with the C9-CRISPR followed by a second transfection with either the long isoform or short isoform. Transfection with C9L isoform led to strong SG formation after dithiothreitol (DTT) exposure; rescuing the inhibition of SG formation following knockdown of all isoforms using the C9-CRISPR. This rescue was not seen for cells transfected with C9S – implying isoform specificity. more precisely that the long isoform is important for SG dynamics.

1.5.5 Summary of C9ORF72 protein function

Prior to the discovery of the repeat expansion located in intron 1 *C9ORF72* in 2011 as a cause of ALS/FTD nothing was known about the proteins function; no previous publications had investigated C9ORF72. Discovering the function of C9ORF72 was originally hampered by the poor quality of commercially available antibodies and difficulties in generating antibodies that could distinguish between the two protein isoforms of C9ORF72. Structural analysis revealed C9ORF72 has homology to the DENN family of proteins suggesting C9ORF72 may have a functional role in endosomal trafficking events, including autophagic and lysosomal pathways. C9ORF72 has been shown to have a functional role in the initiation of macroautophagy and may also regulate lysosomal biogenesis. Both biological pathways are strongly associated with ALS/FTD (Hardy and Rogaeva, 2014). Additionally, mouse studies have revealed the protein appears to regulate immune homeostasis, previously associated with neurodegenerative disease (Heneka et al., 2014) and also SG dynamics (Maharjan et al., 2017); particularly relevant to the pathogenesis of C9ALS/FTD (Lee et al., 2016; Boeynaems et al., 2017a).

1.6 Mechanisms of *C9ORF72* ALS/FTD toxicity

Currently three non-mutually exclusive hypotheses have been proposed as to how expanded G_4C_2 repeats may cause pathogenesis (figure 1.16) (Gitler and Tsuiji, 2016). Firstly, the presence of expanded G_4C_2 repeats within *C9ORF72* leads to reduced transcription of the gene and hence a reduction in C9ORF72 protein levels. The second hypothesis is that the repeats cause disease via RNA-mediated toxicity; through the formation of RNA foci which sequester essential RNA binding proteins. The third proposed mechanism for toxicity is the translation of the expansion from both the sense and anti-sense strand leading to the production of the DPRs which aggregate in patient brain and spinal cord. Defining which of these mechanisms are the driving force behind the disease is crucial to the development of potential treatments; for example, anti-sense oligonucleotides (ASOs) directed against the G_4C_2 RNA would potentially protect against RNA and DPR toxicity but may enhance the effects of a loss of C9ORF72 protein function by lowering C9ORF72 expression (Gitler and Tsuiji, 2016). The evidence for and against these 3 mechanisms will now be discussed in detail.

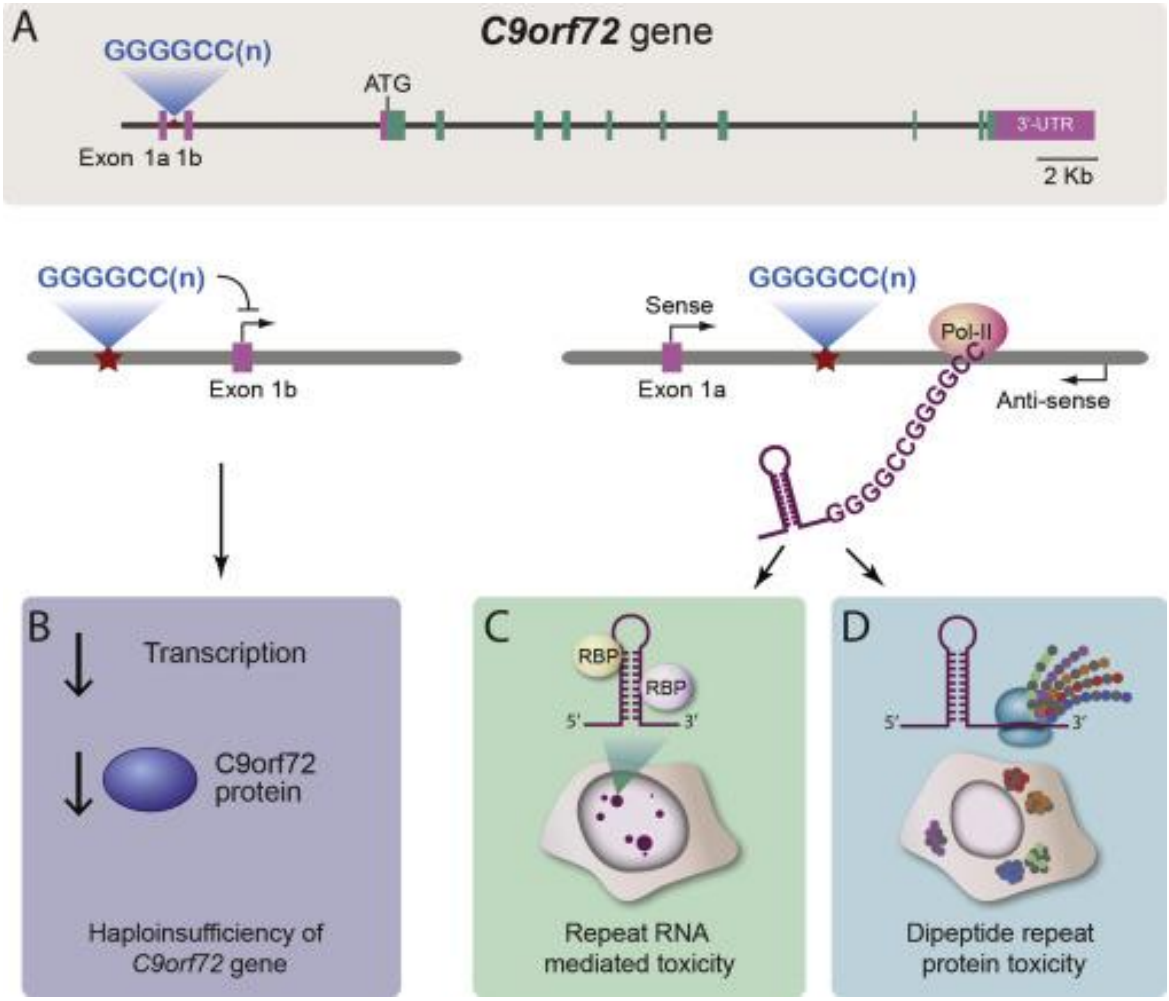

Figure 1.16 | Pathogenic mechanisms associated with the G_4C_2 repeat expansion in *C9ORF72*. (A) The repeat expansion is in intron 1 of *C9ORF72* between non-coding exons 1a and 1b. **(B)** The repeat expansion interferes with the transcription of the gene leading to a reduction in *C9ORF72* gene expression resulting in decreased C9ORF72 protein expression and potentially a loss of function of the protein. **(C)** The sense and anti-sense repeat RNA transcripts form nuclear and cytoplasmic RNA foci which sequester RNA binding proteins involved in numerous RNA processing pathways including splicing leading to defective pre-mRNA splicing. **(D)** The sense and anti-sense transcripts are translated via RAN translation leading to the formation of dipeptide repeat proteins which form inclusions in the brains of ALS/FTD patients which may confer toxicity. From Gitler & Tsuiji, (2016).

1.7 Mechanisms of toxicity: protein loss of function

Analysis of post-mortem C9ALS/FTD brain samples has revealed both transcript (De-Jesus Hernandez et al., 2011; Gijselinck et al., 2012; Belzil et al., 2013; Ciura et al., 2013; Xi et al., 2013; Fratta et al., 2013; Waite et al., 2014; van Blitterswijk et al., 2015) and protein (Xiao et al., 2015a; Sivadasan et al., 2016) levels are decreased. As the C9ORF72 plays a role in trafficking and autophagy then impairments in these cellular functions would be predicted to result from C9ORF72 loss of function. Aoki et al. (2017) reported markedly impaired intracellular and extracellular vesicle trafficking and a defective trans-Golgi network phenotype in C9ALS/FTD iPSC motor neurons; such phenotypes could be rescued by overexpression of C9ORF72 or ASOs designed to upregulate transcript V1. These phenotypes were associated with the interaction of C9ORF72 with RAB7L1. *C9ORF72* repeat expansion associated pathology is associated with the accumulation of unique p62 positive inclusions (Al-Sarraj et al., 2011). The p62 protein is an autophagy receptor known to have an important role in targeting specific cargoes for autophagic degradation (Rusten and Stenmark, 2010). Further a number of genes associated with ALS have roles in autophagy including SQSTM1 (encoding p62), OPTN, VCP, UBQLN2 and TBK1 (Oakes et al., 2017).

Webster et al. (2016a) tested whether this p62 pathology was related to a potential loss of C9ORF72 protein function. A loss of C9ORF72 resulted in the accumulation of p62 positive puncta in both HeLa cells and primary cortical neurons. Reintroducing C9ORF72 reduced the number of p62 puncta to control levels. Webster et al. (2016a) went on to measure basal levels of autophagy in C9ALS/FTD iNeurons. Compared to controls, cell lines from C9ALS/FTD patients displayed higher levels of LC3-I and following treatment with bafilomycin A1, LC3-II levels were also increased but this increase was significantly lower than what was seen in treated control cells. Hence although autophagy was not blocked, basal levels were attenuated in C9ALS/FTD iNeurons compared to controls. Sellier et al. (2016) demonstrated the C9ORF72, SMRC8 and WDR41 complex interacts weakly with p62 and also OPTN, another autophagy receptor. Sellier et al. (2016) also demonstrated lowering C9ORF72 levels alters autophagy, specifically acting upon the formation of autophagosomes. C9ORF72 knockdown led to the formation of p62 aggregates in

primary mouse cortical neurons and GT1-7 neuronal cells. Aggregates of p62 were also observed following SMRC8 and WDR41 knockdown in GT1-7 cells. Re-expressing *C9ORF72* mRNA fully rescued autophagy dysfunction caused by C9ORF72 knockdown; however only the C9L isoform and not the C9S isoform rescued autophagy defects. Importantly decreased C9ORF72 expression promoted the aggregation of TDP-43; shRNA mediated knockdown of Ms-C9orf72 in mouse embryonic cortical neurons led to the accumulation of a ~30kDA fragment of mouse TDP-43 (Tdp-43). Furthermore, using an antibody against phosphorylated serine 409/410 of Tdp-43, cytoplasmic aggregates of Tdp-43 were observed in these neurons (Sellier et al., 2016). In addition, overexpression of C9ORF72 reduced the aggregation propensity of the D196G mutant of TDP-43. The short isoform of C9ORF72 has been implicated in nucleocytoplasmic transport deficits; C9S was found at the nuclear membrane in healthy control spinal motor neurons, however for both C9 and sporadic ALS cases the C9-S isoform was mislocalised to the plasma membrane (Xiao et al., 2015a). The loss of C9-S from the nuclear membrane correlated with a mislocalisation of KPNB1, RAN and TDP-43 in C9ALS spinal motor neurons (Xiao et al., 2015a).

A single zebrafish knockdown model demonstrated loss of the zebrafish homologue of C9ORF72 leads to motor deficits in fish (Ciura et al., 2013). Ciura et al. (2013) used an antisense morpholino oligonucleotide (AMO) in order to block the translation of the 1 zebrafish homologue of C9ORF72 (*z C9orf72*). As with the human gene, the fish homologue also encodes for 3 separate transcript variants that leads to 2 protein isoforms. Two different AMOs were designed and injected into the fish in order to target 2 translation initiation sites contained within the 3 transcripts. Knockdown of *z C9orf72* led to disrupted arborisation and shortening of motor neuron axons compared to controls; further co-injection of the human C9ORF72 long transcript with the 2 AMOs rescued this axonal phenotype. These axonal phenotypes are potentially explained by the role C9ORF72 plays in regulating cellular actin dynamics via its modulation of Arf6 GTPase activity and the phosphorylation of cofilin (Sivadasan et al., 2016). No changes in the levels of *zTDP-43* were observed. C9 AMO injection led to abnormalities in both spontaneous and evoked swimming, these motor deficits could be rescued upon co-expression of *C9ORF72* mRNA. However, another unpublished zebrafish study did not observe any axonal deficits or

impaired motor function in fish following C9ORF72 knockdown (Schmid, Hruscha, Haass, unpublished; as referenced in Stepto et al., 2014). Deletion of the *C. elegans C9orf72* homologue, *F18A1.6*, also called *alfa-1*, leads to age-dependent motor phenotypes eventually causing paralysis and also the specific degeneration of GABAergic motor neurons (Therrien et al., 2013). Furthermore *alfa-1* mutants were highly sensitive to osmotic stress indicating *alfa-1* plays a role in protecting worms from osmotic stress. Therrien et al. (2013) also found the motor impairment caused by *alfa-1* loss was additive with mutant TDP-43, but not FUS, toxicity. Ugolino et al. (2016) found homozygous Ms-C9orf72 knockout mice showed decreased survival compared to controls with 50% being dead by 600 days. A decrease was also observed for heterozygous mice with around 20% being dead after 600 days. The authors note both homozygous and heterozygous mice develop normally before demonstrating rapid progressive lethargy prior to death. As was reported previously splenomegaly was observed in homozygous Ms-C9orf72 null mice (O'Rourke et al., 2016; Atanasio et al., 2016; Burberry et al., 2016; Sudria-Lopezet al., 2016). Loss of Ms-C9orf72 also impaired mTOR signalling in mouse embryonic fibroblasts generated from these mice. Sullivan et al. (2016) observed both autophagy and lysosomal deficits in Ms-C9orf72 deficient mice. Interestingly the mouse homologue of C9ORF72 is enriched in neurons known to be vulnerable and degenerate in ALS and FTD (Suzuki et al., 2013).

1.7.1 Mechanisms of C9ORF72 protein loss of function

Several studies have probed into the mechanisms behind a C9ORF72 protein loss of function. Early studies demonstrated downregulation of C9ORF72 in expansion carriers (DeJesus-Hernandez et al., 2011). In the fontal cortex reduced expression levels have been reported for all *C9ORF72* transcripts (Gijselinck et al., 2012; Ciura et al., 2013; Belzil et al., 2013; Waite et al., 2014), the V1 transcript (Gijselinck et al., 2012) and the V2 transcript (DeJesus-Hernandez et al., 2011; Fratta et al., 2013; Waite et al., 2014). In the cerebellum a decrease in total transcript levels has been reported (Belzil et al., 2013) and a specific reduction in V2 cerebellar transcript levels has also been observed (Waite et al., 2014). The most detailed study investigating *C9ORF72* transcript expression levels has been performed by van Blitterswijk et al. (2015). Using quantitative and digital molecular barcoding techniques the authors

observed a decrease in expression levels in patients compared to controls for total *C9ORF72* transcripts, V1 and V2 levels, with the strongest decrease being seen for V2 - no decreases were seen for transcript V3. Further increased V1 levels in the frontal cortex and cerebellum associated with increased survival. The repeats are located in an intron between two alternatively spliced exons for transcripts V1 and V3; whereas for V2 the expanded repeats are located within the promoter region (Stepto et al., 2014). It has been proposed that this reduced expression of V2 leads to haploinsufficiency if transcription from the wild-type allele is unable to compensate in producing enough C9ORF72 protein (Gitler and Tsuiji et al., 2016) (figure 1.17). Indeed Sareen et al. (2013) observed the presence of G_4C_2 repeats potentially causes transcription of exon 1a (from transcripts V1 and V3) to be favoured over exon 1a (from transcript V2). One mechanism for this downregulation is the methylation of *C9ORF72* discussed previously; this epigenetic change has been intensively studied to determine whether such a change may play a role on C9ORF72 loss of function (Belzil et al., 2016). Hypermethlylation of the 5' CpG island in the *C9ORF72* promotor is found in 10-30% of C9ALS/FTD cases (Belzil et al., 2016). The length of the repeat has been found to influence this hypermethylation; with longer repeats being associated with increases in promoter methylation (Xi et al., 2015b). Xi et al. (2013) demonstrated increased methylation of the 5' CpG island correlates with a shorter duration of disease in ALS. A second mechanism potentially underlying the decreased mRNA expression of *C9ORF72* is abortive transcription due to the presence of the G_4C_2 repeats (Haeusler et al., 2014). Haeusler et al. (2014) found the repeat expansion perturbs normal transcription generating abortive RNA transcripts; this process was facilitated by the formation of G-quadruplexs and RNA: DNA hybrids known as R-loops. These abortive RNA transcripts were bound by specific proteins such as nucleolin, which was mislocalised in patient cells. The binding nucleolin to the repeat RNA was dependent upon the RNA G-quadruplex formed by the abortive RNA transcript.

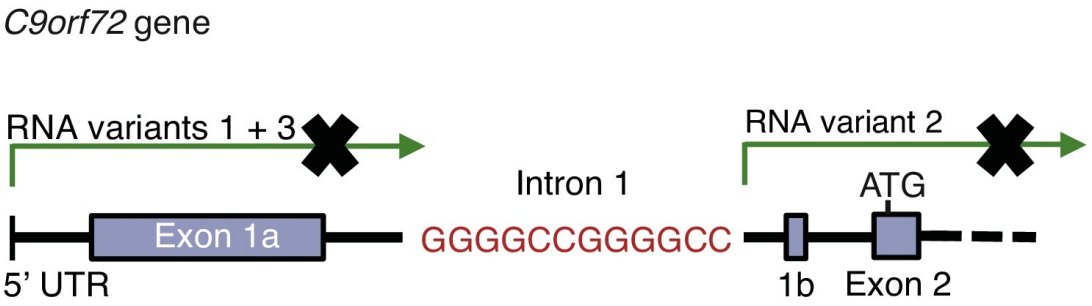

Figure 1.17 | G_4C_2 repeat associated epigenetic silencing associated resulting in reduced *C9ORF72* expression and therefore haploinsufficiency. *C9ORF72* is alternatively spliced to produce 3 mRNA transcripts. The G_4C_2 repeat expansion is located in-between exons 1a and 1b and is proposed to cause an epigenetic silencing of the *C9ORF72* gene through either hypermethylation or abortive transcription leading to reduced C9ORF72 protein expression. Adapted from Moens et al., 2017.

1.7.2 Evidence against loss of protein function

A central role for a loss of C9ORF72 in C9ALS/FTD neurodegeneration is unlikely (Edbauer and Haass, 2016). Despite C9ORF72 knockdown clearly leading to defective autophagy, no reduction in neuronal viability is observed following either siRNA or shRNA treatment of neuronal primary cultures (Sellier et al., 2016). Reduced *C9ORF72* transcript expression because of hypermethylation may be neuroprotective (Belzil et al., 2013; Liu et al., 2014; McMillan et al., 2015; Russ et al., 2015). (Belzil et al., 2016). For example, reduced RNA foci and DPR burden is seen in post-mortem patient brains with CpG hypermethylation (Liu et al., 2014). Bauer, (2016) demonstrated *in vitro* that methylation of the G_4C_2 reduces poly-GP levels and RNA foci formation. Hence reducing C9ORF72 mRNA appears to be neuroprotective as it reduces the toxic the gain of function species derived from the G_4C_2 expansion. Indeed, lowering levels of C9ORF72 in patient iPSC neurons and iPSC motor neurons improves, rather than worsens gene expression changes associated with the repeat expansion (Donnelly et al., 2013; Sareen et al., 2013).

Genetic evidence also suggests a loss of C9ORF72 protein function is not sufficient for the development of clinical disease; no loss of function mutations identified including nonsense or frameshift mutations have been found in *C9ORF72* (Harms et al., 2013). A splice site mutation in *C9ORF72* was found in a Chinese ALS patient (Liu et al., 2016c); however, it is unclear if this mutation is disease causative (Moens et al., 2017). Further a rare homozygous expansion case showed no enhanced pathological or clinical phenotype (Fratta et al 2013). Despite evidence in worms (Therrien et al., 2013) and fish (Ciura et al., 2013) demonstrating a loss of C9ORF72 homologues can cause motor dysfunction and neurodegeneration; this has not been replicated in murine models of C9ORF72 loss of function. ASO knockdown reducing levels of Ms-C9orf72 by up to 50% for several months does not result in any neuropathology or behavioural phenotypes (Lagier-Tourenne et al., 2013). Selective depletion of Ms-C9orf72 from neurons and glial cells also does not produce any motor dysfunction, motor neuron degeneration or reduced survival (Koppers et al., 2015). Furthermore, no TDP-43 mislocalisation or changes ubiquitin staining were seen in these mice. Jiang et al. (2016) observed very mild social interaction defects and a mild reduction in rotarod performance without changes in grip strength, gait,

general activity or a loss of motor neurons in Ms-C9orf72 null mice. Taken together this data strongly suggests a loss of C9ORF72 is not sufficient to cause ALS/FTD phenotypes or neurodegeneration. This indicates a loss of function of the protein is not the primary disease mechanism in C9ALS/FTD.

1.7.3 Summary of loss of protein function

In summary findings from murine knockout models and genetic evidence from patients has shown a loss of C9ORF72 function is unlikely to be the main pathogenic cause of C9ALS/FTD (Edbauer and Haass, 2016). C9ORF72 loss of function does not lead to TDP-43 associated neurodegeneration or any motor neuron cell loss for that matter (Koppers et al., 2015). The role of C9ORF72 in autophagy and immune dysregulation is interesting considering both processes are associated with ALS and FTD (Heneka et al., 2014; Gao et al., 2017). Defective autophagy could enhance DPR accumulation and thereby DPR toxicity (Webster et al., 2016b). Indeed, the DPRs are mainly cleared via autophagy (Cristofani et al., 2017); hence a loss of C9ORF72 protein function may enhance DPR accumulation. Furthermore, the link between C9ORF72 and SG formation (Maharjan et al., 2017) is interesting given the recent findings that the arginine DPRs perturb SG dynamics (Lee et al., 2016; Boeynaems et al., 2017). Hence a loss of C9ORF72 protein function likely plays a modulatory role in disease.

1.8 Mechanisms of toxicity: G_4C_2 RNA toxicity

Analogues to what is seen in other neurodegenerative diseases caused by expanded nucleotide sequences in the non-coding regions, it is hypothesized that the formation of sense G_4C_2 and anti-sense C_4G_2 foci causes the sequestration of RNA binding proteins and splicing factors leading to their loss of function (DeJesus-Hernandez et al., 2011). Indeed, the accumulation of these RNA foci in the C9ALS/FTD brain is similar to the RNA foci identified in myotonic dystrophy type 1 (DM1), myotonic dystrophy type 2 (DM2) and fragile X-associated tremor (FXTAS); these diseases are caused by expanded nucleotide sequences in the non-coding regions of the genes *DMPK*, *CNBP* and *FMR1* respectively (Taylor et al., 2016). In these diseases the RNA foci sequester RNA binding proteins and splicing factors leading to a loss of their normal function which underlies some of the pathogenesis of these diseases (Taylor et al., 2016). For example, in myotonic dystrophy type 1 (DM1) caused by a CTG repeat expansion in the 3'UTR of the *DMPK* gene (Brook et al., 1992); the RNA binding protein muscleblind (MBNL) is sequestered into CUG RNA foci (Fardaei et al., 2001); upregulation of MBNL is sufficient to rescue toxicity *in vivo* (de Haro et al., 2006). Indeed, restoration of MBNL function rescues DM1 phenotypes (Pettersson et al., 2015). Based on this several labs sought to identify RNA binding proteins sequestered by the sense and/or anti-sense foci produced from the *C9ORF72* expansion; and whether a loss of function of these RNA binding proteins contributes to neurodegeneration (Gitler and Tsuiji, 2016).

1.8.1 Sequestration of RNA binding proteins by G_4C_2 RNA

Numerous RNA binding proteins have been found to be sequestered by G_4C_2 or C_4G_2 RNA foci in patient tissue and model systems or bind to the repeat RNA in *in vitro* assays. Interacting proteins include ADARB2, ALYREF, hnRNP A1, hnRNP A3, hnRNP H, hnRNP H1, hnRNP H2, hnRNP F, hnRNP K, hnRNP Q, hnRNP U, Nucleolin, PABPc, Pur-alpha, SRSF1 (SF2) SRSF2 (SC35) and Zfp106 (Donnelly et al., 2013; Lee et al., 2013; Sareen et al., 2013, Almedia et al., 2013, Haeusler et al., 2014; Cooper-Knock et al., 2014; Cooper-Knock et al., 2015; Mori et al., 2016; Rossi et al., 2015; Celona et al., 2017). The RNA binding partners for sense and anti-sense

foci are similar (Haeusler et al., 2014; Cooper-Knock et al., 2015). The RAN GTPase protein RanGAP1 involved in nucleocytoplasmic transport has also been identified as a component of G_4C_2 RNA foci (Zhang et al., 2015). Importantly TDP-43 has not been identified as interacting directly with G_4C_2 RNA foci (Donnelly et al., 2013; Lee et al., 2013; Sareen et al., 2013, Almedia et al., 2013; Lagier-Tourenne et al., 2013; Chew et al., 2015).

The sequestration of RNA-binding proteins into foci (figure 1.18) has been proposed to perturb several aspects of post-transcriptional gene regulation; including alternative mRNA splicing, translational regulation and mRNA transport (Rossi et al., 2015; Barker et al., 2017). Whilst a number of different RNA binding proteins have been found to interact with G_4C_2 or C_4G_2 RNA foci; if any of these are critical for pathogenesis remains an open question (Gitler and Tsuiji, 2016). The binding of proteins to the repeat RNA is structural dependent (Haeusler et al., 2014). The G-rich DNA and RNA of the repeat expansion, due to the highly repetitive nature of the sequence, forms highly complex secondary structures including unimolecular and multimolecular G-quadruplexes (Fratta et al., 2012; Reddy et al., 2013; Haeusler et al., 2014; Zhou et al., 2015; Conlon et al., 2016). G-quadruplexes are stable four strand structures formed of planar guanine tetramers stacked on top of one another (Zhou et al., 2015). Further the G_4C_2 RNA and the anti-sense C_4G_2 RNA both form a stable hairpin secondary structure (Haeusler et al., 2014). Haeusler et al. (2014) demonstrated RNA binding proteins favour specific RNA topologies formed by the repeats; unique proteins bound the G_4C_2 G-quadruplex, the G_4C_2 hairpin and the C_4G_2 hairpin. For example, nucleolin preferentially bound the G_4C_2 G-quadruplex; hnRNP F bound both the G_4C_2 G-quadruplex and hairpin whereas hnRNP K showed preferential binding to the anti-sense C_4G_2 RNA.

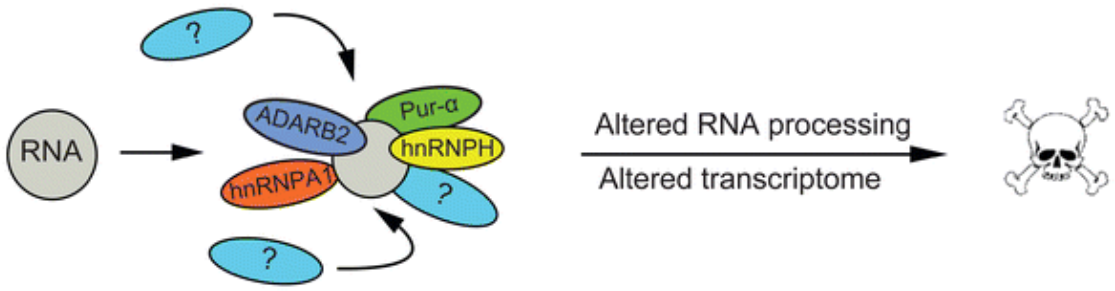

Figure 1.18 | G_4C_2 repeat RNA toxicity. G_4C_2 RNA is potentially neurotoxic by sequestering RNA binding proteins leading to perturbed RNA processing. Adapted from Stepto et al., (2014).

1.8.2 Splicing alterations associated with G_4C_2 RNA toxicity

A number of the proteins that bind to the repeat RNA have important functions in RNA splicing; in particular several members of the hnRNP family of splicing factors (Barker et al., 2017). Lee et al. (2013) found in SH-SY5Y cells' expressing 72 G_4C_2 repeats that splicing efficiency was reduced. hnRNP H is required for the inclusion of exon 7 into the mature *TARBP2* RNA. Lee et al. (2013) observed inclusion of exon 7 into *TARBP2* is dramatically reduced; a lesser reduction in exon 7 inclusion was also seen in cells transfected with the 72 G_4C_2 repeat construct. This data therefore shows the sequestration of RNA binding proteins into RNA foci perturbs splicing and indicates a loss of hnRNP H function may play a role in C9ORF72 ALS/FTD pathogenesis. Indeed Colon et al. (2016) found hnRNP H is the major protein associated with G_4C_2 RNA and 1.9x more insoluble in C9ALS motor cortex compared to controls due to the formation of insoluble hnRNP H aggregates with G-quadruplex RNAs in the C9ALS brain. Colon et al. (2016) also found reduced exon inclusion for a number of known hnRNP H RNA targets from C9ALS cerebellum, most notably ataxin-2 – a disease modifier of C9ALS/FTD (van Blitterswijk et al., 2014). Indeed Colon et al. (2016) note splicing dysregulation, in particular exon skipping, has previously been reported in RNA-seq data from the cerebellum and frontal cortex of C9ALS patients (Prudencio et al., 2015). It was suggested hnRNP H may potentially be driving these splicing changes as misregulated alternative splicing cassette exons in both ALS cerebellum and frontal cortex were found to be enriched

in binding motifs for hnRNP H (Prudencio et al., 2015). Prudencio et al. (2015) found splicing abnormalities were 3-fold higher in the C9ALS brain compared to SALS. In addition, both Cooper-Knock et al., (2015) and Colon et al. (2016) found a reduction in the duration of disease correlated with enhanced alterations in exon inclusion.

1.8.3 Translational deficits associated with G_4C_2 RNA toxicity

Rossi et al. (2015) identified translational repression as one consequence of the sequestration of RNA binding proteins into repeat foci. Rossi et al. (2015) generated a 31 repeat G_4C_2 construct $(G_4C_2)_{31}$ which when transfected in mouse motor neuronal NSC34 cells formed RNA foci. Proteins that bound to the $(G_4C_2)_{31}$ repeats were regulators of translation including initiation and elongation factors (EF1α, eIF2α, eIF2β and eIF2γ) and a number of other proteins that regulate translation including Pur-alpha, Pur-beta, ILF2, ILF3 and RAX. One of these targets Pur-alpha has previously been identified as a binding partner of repeat RNA *in vitro* and in an *in vivo Drosophila* model of repeat toxicity by Xu et al. (2013). Xu et al. (2013) found overexpression of Pur-alpha rescued retinal induced repeat toxicity and was found to form aggregates in post mortem human cerebellum and the *Drosophila* retina. Pur-alpha, although rarely co-localising with RNA foci, was grossly mislocalised in cells expressing the repeats. In un-transfected cells Pur-alpha is diffusely localised in the cytoplasm, however in cells with RNA foci, Pur-alpha, and its binding partner FMRP, accumulated in cytoplasmic and nuclear granules. Rossi et al. (2015) observed increased SG formation (a marker for translational arrest) and a reduction in the rate of mRNA translation in cells transfected with the $(G4C2)_{31}$ construct. Furthermore Rossi et al. (2015) demonstrated the repeats cause defective mRNA export from nucleus, causing poly(A) RNAs to accumulate. Rossi et al. (2015) observed the poly(A) binding protein PABPC1 which has a crucial role in cytoplasmic mRNA translation and stability, accumulated in the nucleus of $(G4C2)_{31}$ cells. Furthermore, PABPC1 accumulated in RNA foci and bound repeat RNA. Barker et al. (2017) suggest nuclear sequestration of PAPBC1 is likely to be detrimental to global translational efficiency of cytoplasmic mRNAs due to its role in promoting the interaction between translation initiation factors and the 5' cap of mRNA, causing an association of the ribosome with the nascent transcript (Barker et al., 2017). Interestingly PAPBC1 inclusions have been found previously in human post-mortem

ALS spinal cord, where intriguingly the colocalization between PAPBC1 and TDP-43 in spinal cord motor neurons was twice as frequent in C9 ALS compared to sporadic ALS (McGurk et al., 2014).

Burguette et al. (2015) demonstrated G_4C_2 repeat RNA perturbs local translation in neurites. The authors detected G_4C_2 RNA particles in neurites that are actively transported to distal neuronal processes, indicting the repeat RNA is assembling into mRNA transport granules. Branching defects were also reported in both spinal cord motor neurons cultured from rats and *Drosophila* sensory dendritic arborisation neurons (Burguette et al., 2015). These branching defects were due to altered transport granule function, when components of mRNA transport granules such as fly fragile x mental retardation protein (*dFMRP*) were up or downregulated the branching phenotypes were potentiated or rescued respectively. FMRP in rat spinal cord neurons was found to colocalise with $(G4C2)_{31}$ in neuritic granules in neuronal processes which suggests FMRP function might be perturbed. FMRP controls the spatial and temporal expression of proteins involved in synaptic plasticity, it is found at postsynaptic dendritic spines where it binds dendritic mRNAs and represses/derepresses their translation (Santoro et al., 2012). Indeed, protein levels of the FMRP target dendritic mRNA postsynaptic density protein (PSD-95); total PSD-95 levels were increased in addition to increases in the number of PSD-95 foci per neuron (Burguette et al. 2015).

These data suggest G_4C_2 RNA disrupts mRNA transport and the translational machinery, leading to transport granule dysfunction. Interestingly it has previously been shown that the G-quadruplex RNA structure is a signal for the targeting of mRNAs to neurites (Subramanian et al. 2011). Taken together the aforementioned studies suggest translational arrest may play an important role in *C9ORF72* repeat expansion mediated toxicity.

1.8.4 Nucleocytoplasmic transport deficits associated with G_4C_2 RNA toxicity

Defective nucleocytoplasmic transport is now established as a key pathogenic mechanism in C9ALS/FTD (Zhang et al., 2015; Freibaum et al., 2015; Boeynaems et al., 2016). Zhang et al. (2015) performed a genetic screen using the 30 G_4C_2 repeat

flies generated by Xu et al. (2013). One of the strongest suppressors of G_4C_2 toxicity was RanGAP, a key regulator of nucleocytoplasmic transport. Human RanGAP was found to bind to the sense G_4C_2 RNA G-quadruplex, additionally RanGAP puncta were also observed in C9ALS iPSCs with the protein also occasionally co-localising with G_4C_2 RNA foci. Cells expressing the 30 G_4C_2 repeat construct showed disturbances in the nuclear/cytoplasmic ratio of Ran, with higher levels of cytoplasmic RAN seen. RanGAP hydrolyses RAN inducing its GTPase activity; catalysing its conversion from RanGTP to RanGDP (Cautain et al., 2014); complex bound GDP RAN binds to cytosolic cargoes and translocates to the nucleus (Stewart, 2007). Hence this data indicates the repeat RNA induces a loss of RanGAP function. Further the nuclear/cytoplasmic ratio of RAN returned to levels of controls when RanGAP was overexpressed or when an ASO that removed RNA foci was used to target the sense G_4C_2 RNA. Nuclear import was also found to be defective in *Drosophila* salivary gland cells expressing 30 G_4C_2 repeats as evidenced by the cytosolic accumulation of a NLS-ΔNES-GFP construct and *Drosophila* TDP-43, *TBPH*. These phenotypes were rescued by the ASO and the G-quadruplex inhibiter TMPyP4 - which would prevent RanGAP interacting with the G-quadruplex G_4C_2 RNA. Based on the above data the authors suggested the nucleocytoplasmic transport deficits were due to *C9ORF72* sense strand RNA toxicity. DPRs were not detectable using dot blotting in the 30 G_4C_2 repeat flies, although the authors were not able to rule out undetectable levels of DPRs potentially having an effect. Indeed, several other studies have shown the DPRs to cause deficits in nucleocytoplasmic transport (Jovičić et al., 2015; Boeynaems et al., 2016a) (see section 1.9.3.1).

1.8.5 Evidence against G_4C_2 RNA toxicity

Animal models provide a conflicting picture as to whether the G_4C_2 RNA plays a toxic role in disease. Xu et al. (2013) generated a 30 sense repeat fly model, the toxicity of which was attributed to the G_4C_2 RNA by Zhang et al. (2015) as DPRs were not detectable under most conditions. Several other *Drosophila* studies however argue against RNA toxicity as a driving factor behind pathogenesis (Mizielinska et al., 2014; Tran et al., 2015). Mizielinska et al. (2014) generated RNA-only G_4C_2 repeat constructs by inserting one of three 6-base-pair interruptions, each of which contained a STOP codon in both the sense and antisense direction every 12 G_4C_2

repeats, which meant there was a stop codon present for all 6 reading frames. These RNA-only constructs developed were equivalent in length to 36, 107 and 288 pure G_4C_2 repeats. The RNA-only constructs were shown to still form the characteristic RNA G-quadruplex formed by pure G_4C_2 RNA repeats. The RNA-only constructs also formed RNA foci at the same propensity as pure repeat constructs of an equivalent length. To compare RNA toxicity to DPR toxicity Mizielinska et al. (2014) generated transgenic flies harbouring these RNA-only constructs, flies with pure G_4C_2 repeats and DPR only flies. DPR only constructs were generated using alternative non-G_4C_2 codons that produce for the same peptide from an ATG. Alternative codon DPR only flies were generated for all DPRs except poly-GP. Expression of either 36 or 103 pure repeats in the *Drosophila* eye (in which RAN translation occurred) caused degeneration, however RNA-only constructs of an equivalent length caused no degeneration, and thus pure repeat toxicity is attributable to DPRs rather than G_4C_2 RNA. When DPR producing pure repeat constructs were expressed in adult fly neurons, a reduction in life-span was seen, whereas RNA only repeats had no effect. Additionally, treating flies with cycloheximide which reduces protein synthesis rescued the lifespan deficits seen in the pure repeat flies; demonstrating toxicity in the pure repeat flies is likely due to DPR expression. Furthermore poly-GR and poly-PR protein only constructs caused both severe neurodegeneration and markedly reduced in lifespan in flies. Tran et al. (2015) generated a *Drosophila* model in which the G_4C_2 repeats were in an intronic context. Flies were generated containing a minigene with exon 1, part of intron 1 with the G_4C_2 repeats and exon 3 of *C9ORF72*. This differs from the Mizielinska et al. (2014) flies and other Drosophila models (Xu et al., 2013; Zhang et al., 2015; Freibaum et al., 2015) in which the G_4C_2 repeats are transcribed as part of a poly(A) mRNA. Flies expressing 106 intronic G_4C_2 repeats develop normally and had no alterations in lifespan. Numerous sense RNA foci were seen, however, rRNA biogenesis and mRNA processing were both unperturbed. Only a small number of differentially expressed genes and exon skipping events were seen suggesting that if RNA foci are sequestering RNA binding proteins in these flies, it is having little effect on RNA processing. When the 106 intronic repeat flies were compared to the 36 repeat pure G_4C_2 flies generated by Mizielinska et al. (2014) in which the repeats are part of the poly(A) mRNA (in which eye degeneration and life span abnormalities are seen), DPR levels were found to be >100 fold higher in the poly(A) 36 repeat flies.

Modest reductions in lifespan could be observed in the 106 intronic repeat flies when they were raised at 29°C; however, DPR production was 4x higher under these conditions compared to basal levels. Hence this study further supports the conclusions of Mizielinska et al. (2014) indicating toxicity is primarily driven through DPRs, rather than sense RNA foci.

BAC transgenic mouse models also argue against RNA toxicity. BAC models allow repeat sizes found in the human cases to be modelled *in vivo*. Peters et al. (2015) and O'Rourke et al. (2015) both generated BAC transgenic mice harbouring expansion sizes pathogenic in humans (500 and ~100-1000 repeats respectively). These mice displayed the typical pathological features seen in patient brains including widespread sense and anti-sense RNA foci in brain regions such as the frontal cortex, primary motor cortex, cerebellar neurons and spinal cord motor neurons. Despite wide-spread RNA foci formation no neurodegeneration, cognitive or motor phenotypes were seen, additionally these mice had a lifespan similar to controls. DPRs were also detected in these models although only poly-GP was probed for in both studies. Liu et al. (2016a) generated numerous BAC transgenic mice of different repeat lengths that displayed motor phenotypes, cognitive decline, muscle denervation, a loss of motor neurons and cortical degeneration. Sense and anti-sense foci were detected in numerous brain regions in addition to aggregates of all three sense DPRs, poly-GP, poly-GA and poly-GR in addition to Tdp-43 pathology. The anti-sense foci in particular were found to accumulate in vulnerable brain regions in these mice. In one of the BAC lines harbouring a short expansion (36 and 29 repeats) DPRs were produced but no RNA foci were seen. These mice demonstrated motor phenotypes, cognitive impairment and end stage Tdp-43 pathology arguing against RNA foci toxicity. Chew et al. (2015) presented evidence in favour of sense G_4C_2 RNA toxicity. Using the adeno-associated virus AAV2/9 to express 66 G_4C_2 repeats in the nervous system of mice Chew et al. (2015) demonstrated nuclear sense RNA foci, all 3 sense DPRs, Tdp-43 pathology, neurodegeneration, motor and cognitive impairment. Interestingly Tdp-43 pathology in this model correlated with the presence of nuclear RNA foci; suggesting repeat RNA may drive TDP-43 dysfunction. It has also been shown the expression of sense repeats from an intron causes toxicity in mouse primary cortical and motor neurons without detectable DPRs (Wen et al., 2014).

However, a recent in-depth clinico-pathological study of RNA foci in a large cohort of C9ORF72 expansion carries suggest RNA foci are not associated with clinico-pathological variability seen in expansion carries nor does it follow the pattern of neurodegeneration (DeJesus-Hernandez et al., 2017). In fact, increased anti-sense RNA foci burden was associated with a delayed age of onset. Sense RNA foci were not associated with age of onset and neither sense nor anti-sense foci were associated with repeat length, C9ORF72 transcripts, DPR proteins, gender, disease subgroups or survival after onset. For example, increased foci burden was not seen in C9FTD frontal cortex compared to C9ALS frontal cortex, which would be predicted if foci are toxic. Most RNA foci were present in cerebellar Purkinje neurons; however, no loss of Purkinjie neurons has been reported in the cerebellum of C9ALS/FTD patients (DeJesus-Hernandez et al.., 2017; Tran et al., 2016). It has been suggested RNA foci may act as neutral intermediates and may even be neuroprotective by preventing the export of the repeat RNA into the cytoplasm where it is translated into DPRs (Tran et al., 2015).

1.8.6 Summary of RNA toxicity

There is clear evidence to demonstrate the formation of RNA foci disrupts numerous cellular processes including splicing, translation and nucleocytoplasmic transport. The findings from *in vivo* models and studies of foci burden in the human brain and their correlations with disease variables suggest the repeat RNA is at least not the sole contributor to C9ALS/FTD pathogenesis; with some going as far as proposing the repeat RNA may even be neuroprotective (Tran et al., 2015). One issue with the *in vivo* models is that although these animals express the repeat RNA they may not have the key RNA binding proteins whose sequestration could be important for RNA toxicity, especially in motor neurons (Liu et al., 2016c). Disentangling the contribution of sense and anti-sense transcripts will also be important for further understanding the contribution of RNA toxicity to disease. At present no study has investigated the potential toxicity of anti-sense C_4G_2 RNA in isolation; indirect evidence suggests the anti-sense foci may have bigger impact in disease associated cell loss (Cooper-Knock et al., 2015; Liu et al., 2016c).

1.9 Mechanisms of toxicity: DPR toxicity

In order to investigate DPR toxicity in isolation from the G_4C_2 RNA researchers have taken advantage of the degeneracy in the genetic code. By using alternative non G_4C_2 triplet codons that code for the same amino acids that constitute the 5 DPRs translated from the different G_4C_2 reading frames, constructs have been generated in which the induvial DPR proteins are produced from an ATG. (Mizielinska et al., 2014; Zhang et al., 2014; May et al., 2014; Tao et al., 2015; Yamakawa et al., 2015; Yang et al., 2015; Zhang et al., 2016; Callister et al., 2016; Boeynaems et al., 2016a; Lee et al., 2016; Lee et al., 2017; Schludi et al., 2017). This allows the potential toxicity of the different DPRs to be studied in isolation from the repeat RNA and also the other DPRs, as RAN translation will not occur. One issue with such models is that they lead to a massive overexpression of the DPRs due to the use of ATG, which may overestimate their contribution to pathogenesis, as RAN translation is less efficient than canonical translation (Kearse et al., 2016). Being able to potentially turn off and on RAN translation would be the most accurate way to study and differentiate between repeat RNA and DPR toxicity and also accurately assess the relative contribution of the DPRs to pathogenesis as the amount of DPR protein will be similar to physiological patient levels. Such work requires a better understanding about the mechanisms of RAN translation and how it initiates.

1.9.1 Mechanisms of RAN translation

Repeat associated non-ATG (RAN) translation is the production of peptides from an mRNA sequence lacking a traditional ATG start codon (Green et al., 2016). RAN translation was first discovered by Zu et al. (2011) in their study of spinocerebellar ataxia type 8 (SCA8). Zu et al. (2011) found when studying the poly-glutamine (poly-Gln) produced from *ATXN8* expansion transcripts, that mutation of the only ATG codon 5’ of the CAG expansion surprisingly did not prevent the translation of poly-Gln. Further translation of the CAG repeat was found to initiate in multiple reading frames generating a total of 3 homopolymeric proteins, poly-glutamine (poly-Gln), poly-alanine (poly-Ala) and poly-serine (poly-Ser). The authors found RAN translation to be dependent on the secondary structure of the CAG RNA as reducing either the number of CAG repeats or the GC content inhibited RAN translation. Zu et

al. (2011) went on to generate antibodies against the poly-Ala peptide finding the protein is present in the cerebellum of SCA8 patients and mouse models of SCA8. Further *in vitro* evidence was presented for poly-Gln anti-sense translation from *DMPK* antisense transcripts with CAG expansion linked myotinic dystrophy type 1 (DM1). Since this initial discovery, RAN translation was been found to occur in a number of neurological diseases associated with nucleotide repeat expansions including C9ALS/FTD (Mori et al., 2013a; Ash et al., 2013), fragile X tremor ataxia syndrome (FXTAS) (Todd et al., 2013), Huntington disease (HD) (Bañez-Coronel et al., 2015), spinocerebellar ataxia type 31 (SCA31) (Ishiguro et al., 2017) and most recently in myotonic dystrophy type 2 (DM2) (Zu et al., 2017). There is increasing evidence that these RAN translated proteins are not just pathological bystanders but rather, actively contribute to the pathogenesis of disease (Green et al., 2016).

How RAN translation occurs and the mechanisms behind this non-canonical form of translation are not yet fully understood. Much of the current understanding into mechanisms of RAN translation has come from work studying the CGG repeat expansion in the 5'UTR of *FMR1* associated with FXTAS (Todd et al., 2013; Kearse et al., 2016). Todd et al. (2013) found RAN translation of CGG repeats occurs in at least 2 reading frames; the GGC (+1) frame producing a polyglycine peptide (FMRpolyG), and the GCG (+2) frame leading to the production of a polyalanine peptide (FMRpolyA). Todd et al. (2013) found both the sequence 5' of the repeats and repeat length are important factors in influencing the levels of RAN translation. For example, it was observed that the insertion of a STOP codon upstream of the CGG repeat sequence inhibited FMRpolyG production. Additionally, mutation analysis revealed FMRpolyG translation occurs at multiple near cognate AUG start codons in the 5'UTR. STOP codons however did not prevent translation of FMRpolyA indicating translation occurs within the repeats itself for this reading frame. Hence RAN translation appears to be mechanistically different in different reading frames (Green et al., 2016). Further research by the same group revealed RAN translation in FXTAS is cap-dependent in multiple reading frames, and does not utilize cap-intendent mechanisms such as an internal ribosomal entry sites (IRES) (Kearse et al., 2016). Canonical translation usually beings with recognition of the 5' methyl-7-guanosine (m^7G) cap on mRNA by the eIF4F complex (Green et al., 2016). Kearse et al. (2016) demonstrated when the m^7G cap is substituted with an A-cap,

not recognized by eIF4, RAN translation from CGG repeats is inhibited. Further by sequestering eIF4E using excess free m^7G cap, RAN translation of both FMRpolyG and FMRpolyA is inhibited. Neither of these modifications disrupted IRES translation, indicating CGG RAN translation in the *FMR1* 5'UTR is cap-dependent similar to the initiating stage of canonical translation (Kearse et al., 2016; Green et al., 2016). Following cap-dependent initiation the 43S pre-initiation complex (PIC) joins the eIF4F complex still bound to the m^7G cap and scans through the 5' UTR in search of an AUG codon (Green et al., 2016). Kearse et al. (2016) note that if this scanning model holds for CGG expansion in FXTAS, then the 43S PIC would have to scan through the CGG repeats in the 5'UTR until it reaches the AUG of FMR1. Indeed, the authors found by inhibiting 43S PIC scanning RAN translation from both CGG reading frames was attenuated – indicating RAN translation requires ribosomal scanning.

Both Todd et al. (2013) and Kearse et al. (2016) have a proposed a model in which the secondary structure of the CGG repeats causes the scanning ribosome to stall increasing the chances of translation occurring at AUG-like codons in sequence optimal positions. Green et al. (2016) suggest a stalling scanning ribosome causes congestion of 43S PICs on the mRNA sequence upstream of the repeat and increases the time the 40S subunit spends on AUG-like codons. Furthermore, Green et al. (2016) propose stalling is likely to favour the dissociation of important eIFs required for AUG start codon recognition or cause changes in ribosomal confirmation similar to what occurs in IRES translation. Green et al. (2016) go on further to argue such occurrences will likely increase the speed of enzymatic catalysis and 48S complex formation at AUG-like codons upstream of the repeats. Green et al. (2016) propose a working model of CGG RAN translation in FMR1 in which both the eIF4F complex and 43S PIC bind to the m^7G cap on FMR1 mRNA; this complex scans through the 5'UTR where it stalls upon encountering the secondary structure of the CGG repeats or the surrounding 5'UTR sequence. This ribosomal stalling leads to the initiation of translation at AUG-like codons either upstream or within the repeats leading to FMRpolyG and FMRpolyA production.

This model has been supported by a recent study by Sellier et al. (2017). Sellier et al. (2017) using a novel CGG mouse model found RAN translation is dependent on a ACG near-cognate start codon found within a putative kozak sequence providing an

optimal context for the initiation of translation. Additionally the N terminus of FMRpolyG starts with a methionine, strongly indicating canonical translation is occurring. Furthermore, translation initiated at the ACG codon was independent of the CGG repeats, as small FMRpolyG peptides were detected when fused to GFP in controls and even without any CGG repeats. Sellier et al. (2017) note this indicates the existence of a short upstream open reading frame (uORF); such uORFs are normally translated into small undetectable peptides but are detectable when fused with large tags due to the formation of a stable protein. However, an expansion of over 70 CGG repeats was required to generate a large enough uORF to create a stable, detectable FMRpolyG protein. This data is consistent with ribosomal stalling at CGG hairpins promoting initiation of translation at near-cognate start codons 5' of the repeats. Hence these results indicate FMRpolyG is not a product on RAN translation; rather it is translated via conventional ribosome scanning at an upstream ACG codon decoded by Met-tRNA; the ribosomal complex reads through 55-200 CGG codons and before finally terminating at a TAA STOP codon within a differing reading frame to the FMR1 coding region. Translation of the FMRpolyA peptide was not investigated, with previous work suggesting its translation is a result of RAN translation with initiation beginning in the repeats (Todd et al., 2013).

Mechanisms of RAN translation of the G_4C_2 expansion is currently unknown and no studies have as of yet looked into this. Some insight has come from G_4C_2 constructs used to study repeat toxicity. For example, when the G_4C_2 repeats are placed in the 5'UTR, RAN translation of DPRs is clearly repeat length dependent; with larger repeat sequences producing more DPR protein (Mori et al., 2013a; Zu et al., 2013; Su et al., 2014). For example, Jiang et al. (2016) developed BAC transgenic mice with either 110 repeats or 450 repeats; despite RNA levels being higher in the 110 repeat mice, RAN translation was only seen in mice harbouring 450 repeats. Green et al. (2016) note however that this repeat length requirement appears to be different for different reading frames and in different sequence contexts. For example, all DPRs are detectable at repeat unit lengths ranging between 30 – 40 repeats placed downstream of a synthetic sequence (Zu et al., 2013). Mori et al. (2013a) found when the G_4C_2 repeats are placed 113 nucleotides downstream of intron 1; that the repeat length required for RAN translation within this sequence context differed for the different sense DPRs. For example, poly-GA was detected at 38 repeats, poly-

GP at 66 repeats and poly-GR was not detectable up to 145 repeats. Gendron et al. (2013) also reported a length requirement for the anti-sense DPRs, with poly-GP and poly-PR detectable from 66 repeat length construct whilst poly-PA was not observed. Sellier et al. (2017) note that the G_4C_2 repeats which are 5' to the *C9ORF72* ORF are in frame with a near-cognate CTG codon within a correct kozak sequence (gctCTGg) encoding poly-GA; which incidentally is the most abundant DPR species (Schludi et al., 2015). Examination of the flanking sequences around the anti-sense C_4G_2 transcript shows the presence of an ATG codon in both the poly-PR and poly-GP reading frames (Zu et al., 2013); hence a mixture of canonical translation and RAN translation may occur for these anti-sense DPRs (Cleary and Ranum., 2017). If true this would fit in with the previously discussed finding that poly-GP from the anti-sense strand is more common than poly-GP from the sense strand (Zu et al., 2013).

Unlike the CGG repeats in FXTAS which are found within the 5'UTR of *FRM1*, the G_4C_2 repeats are located in an intron so should be spliced out and not come in contact with the scanning ribosome; hence at present it is unclear what RNA species is responsible for RAN translation of the G_4C_2 repeats (Green et al., 2016). Green et al. (2016) note there are three potential RNA species that may act as the template for G_4C_2 RAN translation; a retained intron, a spliced intron in a lariat, or an aberrant disease-specific transcript formed from stalling during transcription. Indeed, how the repeat containing intron is processed remains an important question in the field (Barker et al., 2017). Experimental evidence supports the first of these potential templates; it has been demonstrated that intron 1 containing the G_4C_2 repeats are retained in a proportion of polyadenylated *C9ORF72* mRNA transcripts and exported to the cytoplasm where they are thought to be subject to RAN translation (Niblock et al., 2016; Hautbergue et al., 2017). A larger number of transcripts containing intron 1 have been detected in *C9ORF72* patient brain tissue compared to controls (Mori et al., 2013b; Niblock et al., 2016) indicating a preferential usage of exon 1a due to the presence of the repeats, which would mean increases in RNA levels of transcript containing variants (Barker et al., 2017). Splicing of *C9ORF72* intron 1 seems not to be altered by G_4C_2 repeats as the amount of unspliced *C9ORF72* transcripts as measured by the exon1-intron 1 junction is unchanged between control and patient tissue (Tran et al., 2015). The mechanisms behind this intron 1 retention is unknown however intron 1 retention was seen in *C9ORF72* polyadenylated RNA from both

wild-type and alleles containing the expansion suggesting intronic retention of intron 1 is part of the normal processing of *C9ORF72* transcripts (Niblock et al., 2016). The majority of intron 1 retaining C9ORF72 transcripts were found to accumulate within the nucleus with a small proportion detected in the cytoplasm (Niblock et al., 2016).

Based on these findings it has been proposed by Niblock et al. (2016) that the nuclear intron 1 retaining *C9ORF72* mRNA is not degraded due to the presence of the G_4C_2 repeats, hence aggregating forming nuclear RNA foci. Some of the *C9ORF72* mRNA containing intron 1 is however exported to the cytoplasm where it is subject to RAN translation producing the DPRs. Indeed, indirect support of this hypothesis comes from Hautbergue et al. (2017) who found knockdown of the nuclear export adaptor SRSF1 rescues neurodegeneration in *Drosophila* and patient derived motor neurons. SRSF1 depletion prevents nuclear export of *C9ORF72* mRNA and reduces DPR levels. Interestingly the nuclear export of *C9ORF72* transcripts without the expansion was not dependent on SRSF1. van Blitterswijk et al. (2015) found both intron 1 containing (1a and 1b) transcripts correlated positively with levels of poly-GP and poly-GA in the cerebellum of expansion carriers. No association was found between intron 1 containing transcripts and poly-GP levels in the frontal cortex. van Blitterswijk et al. (2015) also observed an increase in the expression levels of transcripts harbouring sequences 5' of the repeats (intron 1a) but not for transcripts 3' of the repeats (intron 1b) in the frontal cortex, but not cerebellum, providing evidence to suggest at least some of the intron 1 containing transcripts are truncated. The authors suggest such truncated transcripts may not be suitable templates for RAN translation, and that the cerebellum may contain less truncated transcripts. Indeed, no increase in intron 1a containing transcripts was observed for this brain region and as just noted poly-GP and poly-GA levels in the cerebellum are associated with levels of both intron 1a and 1b transcripts, with this correlation being most pronounced for intron 1b, indicating the entire first intron is the main template for RAN translation - at least in the cerebellum. Further, antibodies used to detect the DPRs directed against the sequences flanking the repeats efficiently detect the RAN peptides (Zu et al., 2013). Tran et al. (2015) observed RAN translation in *Drosophila* is markedly reduced when the G_4C_2 repeats are located within an efficiently spliced intron arguing against a spliced intron in a lariat being used as the major template for RAN translation (Green

et al., 2016). Green et al. (2016) note however that DPRs are still seen in this model when the flies are raised at higher temperatures; indicating an intron lariat is still potentially capable of being a template for RAN translation.

Proteins sequestered by the G_4C_2 RNA foci may also influence RAN translation. It has recently been shown by Zu et al. (2017) that RAN translation of the CCTG·CAGG expansion in DM2 to produce poly-LPAC and poly-QAGR tetrapeptide RAN proteins is modulated by the sequestration and nuclear retention of the repeat RNA by muscleblind like splicing regulator 1 (MBNL1). RAN translation could be blocked by MBLN1 overexpression potentially via sequestration of the repeat RNA into the nucleoplasm; whereas reducing levels of MBLN1 had the opposite effect. Further in cells RNA foci were inversely correlated with RAN protein accumulation. Whether such a cross-talk between G_4C_2 RNA gain of function effects and RAN translation in C9ALS/FTD exists remains an interesting avenue to explore. Interestingly, wt-TDP-43, FUS and hnRNPA2B1 were shown to modulate RAN translation in SCA31 by potentially acting as chaperones for repeat RNA; thereby potentially preventing the export of the RNA into the cytoplasm (Ishiguro et al., 2017). It will be important for future studies to perform genetic screens to look for proteins that specifically modulate RAN translation of the G_4C_2 transcript and RAN translation in general.

1.9.2 Poly-GA toxicity

Poly-GA is the most abundant of the 5 DPR species in the human brain (Schludi et al., 2015). Poly-GA has strong aggregation properties due to the biophysical properties of the peptide (Freibaum and Taylor, 2017), the peptide is composed of small hydrophobic residues which are uncharged; it is expected to collapse into poorly soluble globules with strong aggregation tendencies (Lee et al., 2016; Freibaum and Taylor, 2017). Indeed poly-GA forms amyloidogenic fibrils which are positively stained by Congo red and thioflavin T (Chang et al., 2016; Freibaum and Taylor., 2017) and is predominantly found in the urea soluble fraction when expressed in cells (Lee et al., 2016). The poly-GA amyloidogenic fibrils form a parallel β-sheet structure with similar structural properties to the amyloid-beta protein in Alzheimer's disease (Chang et al., 2016; Freibaum and Taylor, 2017). Poly-GA has shown to be toxic in many *in vitro* and *in vivo* model systems (Zhang et al., 2014;

May et al., 2014; Mizielinska et al., 2014; Yamakawa et al., 2015; Zhang et al., 2016; Schludi et al., 2017; Khosravi et al., 2017; Lee et al., 2017).

1.9.2.1 UPS dysfunction and ER stress associated with poly-GA

Initial studies using poly-GA only alternative codon constructs by both Zhang et al. (2014) and May et al. (2014) showed poly-GA expression is cable of inducing cellular toxicity. In both papers poly-GA was highly aggregate prone in transfected HEK293T cells, forming both cytoplasmic and nuclear aggregates. Poly-GA inclusions co-localised with ubiquitin and p62 and were dot-like and star-shaped and comprised of filamentous structures; features reminiscent of the poly-GA aggregates seen in FTLD/ALS C9+ patients. Interestingly both papers used HEK293T cells, yet only Zhang et al. (2014) saw toxicity. However, poly-GA expressed in cortical and hippocampal neurons was toxic and rapidly formed aggregates co-localising with both p62 and ubiquitin. Furthermore, neurons had reduced neurite outgrowth and increased caspase3 activation. May et al. (2014) saw no co-localization of poly-GA with TDP-43 – nor were TDP-43 protein levels altered; such observations fit with patient data with a lack of spatial correlation between DPR pathology and TDP-43 pathology (Mackenzie et al., 2015). Interestingly May et al. (2014) found the poly-GA inclusions in their cell culture model were more numerous, but smaller and the GFP signal less intense when compared to poly-GA aggregates in C9+ patient neurons, leading them to suggest they may underestimate poly-GA toxicity. The formation of ubiquitinated poly-GA aggregates suggests an impairment of the ubiquitin proteasome system (UPS); a notion converged on in both papers. May et al. (2014) performed mass spectrometry with most top hits being members of the UPS. The paper focuses on Unc119, a myristoyl-binding protein, which they go on to show co-localises with GA inclusions in both rat primary neurons and post-mortem patient tissue. Zhang et al. (2014) note a consequence of UPS dysfunction would be endoplasmic reticulum (ER) stress. Indeed, levels of numerous ER stress markers in the PERK-CHOP pathway were enhanced in poly-GA transfected neurons and patient tissue. Furthermore, ER stress inhibitors salubrinal and TUDCA rescued poly-GA-induced toxicity. Schludi et al. (2015) found abundant Unc119 pathology in the C9ALS/FTD frontal cortex and cerebellum that resembled poly-GA DPR

pathology; in addition, poly-GA and Unc119 pathology was more severe in the C9FTD cerebellum compared to C9ALS or C9ALS/FTD cerebellar tissue.

Poly-GA is also toxic *in vivo*; Zhang et al. (2016) generated a poly-GA only mouse model. These poly-GA mice developed brain atrophy with neurodegeneration seen in the cortex and hippocampus in addition to astrogliosis. Further poly-GA mice demonstrated anxiety-like behaviour as well as motor and cognitive deficits. The UPS-related proteins HR23A and HR23B both formed nuclear and cytosolic inclusions in the brains of poly-GA mice and were sequestered by poly-GA inclusions further indicting poly-GA exerts toxicity via UPS dysfunction (Zhang et al., 2014; May et al., 2014; Zhang et al., 2016). Zhang et al. (2016) also observed HR23 inclusions in the mice expressing 66 G_4C_2 repeats developed by Chew et al. (2015) and in post-mortem human hippocampus where the inclusions again co-localised with poly-GA. Transfection of HEK293T cells with poly-GA led to HR23A and HR23B cytosolic accumulation with concomitant nuclear depletion. HR23 stabilises the DNA-binding protein xeroderma pigmentosum C (XPC). Poly-GA expression led to enhanced XPC degradation as evidenced by reduced XPC levels in poly-GA mice compared to controls in addition to the presence of XPC inclusions that co-localised with poly-GA in cortical and hippocampal neurons. Hence the co-aggregation of poly-GA and HR23 appears to lead to a loss of HR23 function. Furthermore, co-immunoprecipitation revealed only poly-GA and not poly-GP or poly-GR interacted with HR23B; and analysis of the brains of 66 repeat G_4C_2 mice revealed 85% of HR32B inclusions were positive for poly-GA indicating poly-GA and not the other sense DPRs is the primary cause of H32 dysfunction. Zhang et al. (2016) also demonstrated poly-GA sequestered the nuclear pore proteins RanGAP1 and Pom121 indicating poly-GA also disrupts the nuclear pore complex; the authors note this finding may partially contribute to poly-GA induced HR23 protein mislocalisation.

The toxicity of poly-GA in murine models has been replicated by Schludi et al. (2017). In their model mice developed poly-GA inclusions primarily in motor neurons, the interneurons of the spinal cord and brain stem and deep cerebellar nuclei. However, poly-GA did not lead to any overt neuronal loss in these mice; despite this the authors also observed poly-GA sequestered the UPS-related protein Rad23b; although no sequestration of Unc119 or RanGAP1 as has previously been reported was observed. The Hsp-70 associated protein, Mlf2, co-aggregated with poly-GA in

the spinal cord of mice; Mlf2 pathology was also seen in patient post-mortem brain tissue. Unlike Zhang et al. (2016) who observed astogliosis but not microgliosis, Schludi et al. (2017) saw no astroglisos in their mice but did see marked microglial activation. Indeed, markers for neuroinflammation were upregulated in the spinal cord of 6 month old poly-GA mice. Schludi et al. (2017) note microglial activation correlates with the progression of disease in ALS, and *C9ORF72* patients demonstrate increased levels of microglial activation as compared to non *C9ORF72* ALS patients (Brettschneider et al., 2012). A progressive decline in motor performance was also observed in these mice. Zebrafish (Ohki et al., 2016) and *Drosophila* (Mizielinska et al., 2014) models of poly-GA also show the poly-GA peptide to be toxic. Poly-GA was highly toxic to fish although toxicity did not correlate with aggregation as poly-GA inclusions were exclusively found in the musculature fish; however overall muscle structure was not affected (Ohki et al., 2016). In flies poly-GA leads to a modest reduction in lifespan compared to controls (Mizielinska et al., 2014).

1.9.2.2 Summary of poly-GA toxicity

It has been argued that as poly-GA inclusions are by far the most abundant DPR species, that poly-GA has a critical role in pathogenesis (Schludi et al., 2015). However, although poly-GA seems to be toxic when expressed at high levels (Zhang et al., 2014; May et al., 2014; Yamakawa et al., 2015; Ohki et al., 2016; Zhang et al., 2016; Lee et al., 2017) when compared head to head with the arginine DPRs in *Drosophila*, poly-GA is less toxic (Freibaum and Taylor, 2017; Mizielinska et al., 2014; Wen et al., 2014; Yang et al., 2015; Freibaum et al., 2015; Lee et al., 2016). Freibaum and Taylor, (2017) suggest this may be due to the lower expression levels of poly-GA in these studies compared to those using viral-mediated expression systems; hence it remains to be seen whether poly-GA is toxic at physiologically relevant levels (Freibaum & Taylor, 2017).

1.9.3 Poly-GR and Poly-PR toxicity

The arginine rich DPRs, poly-GR and poly-PR, are highly positively charged and polar due to the presence of arginine residues (Freibaum and Taylor, 2017). Both are predicted to have a flexible coil structure (Freibaum and Taylor, 2017). Biochemical characterisation of a $(GR)_3$ peptide by Flores et al. (2016) revealed poly-GR forms spherical aggregates as visualised by transmission electron microscopy. Secondary structure predicted from circular dichroism spectra indicates 55% of the protein has turn and/or random coil conformation and the other 41% of the protein is β-sheet (Flores et al., 2016). Initial studies in *Drosophila* using alternative codon ATG constructs revealed poly-GR and poly-PR to be highly toxic, causing severe eye degeneration and a dramatic reduction in lifespan (Mizielinska et al., 2014). Both DPRs are much more potent toxins in *Drosophila* when compared with the effects of the other DPRs; with poly-GA the only other DPR to show mild toxicity (Mizielinska et al., 2014). The high toxicity of poly-GR and poly-PR has been replicated in numerous other *Drosophila* studies (Wen et al., 2014; Freibaum et al., 2015; Yang et al., 2015; Lee et al., 2016) and also cellular models and human iPSCs (Kwon et al., 2014; Wen et al., 2014; Tao et al., 2015; Lee et al., 2016). Poly-PR is far more stable than poly-GR; with a half-life of ~72h compared to <30min for poly-GR (Kwon et al., 2014). The mechanisms of poly-GR and poly-PR induced toxicity have been revealed over recent years; these will now be discussed.

1.9.3.1 Poly-GR and poly-PR associated nucleocytoplasmic transport deficits

As discussed in section 1.8.4, defective nucleocytoplasmic transport has been identified as a pathogenic mechanism in C9ALS/FTD (Zhang et al., 2015). Zhang et al. (2015) attributed defective nucleocytoplasmic transport to the G_4C_2 RNA rather than DPRs. However, under certain conditions the flies used in this study can produce poly-GR, hence it cannot be ruled out that undetectable levels of poly-GR contribute to the defective nucleocytoplasmic transport observed in this model. A second *Drosophila* study published in parallel by Freibaum et al. (2015) also reported nucleocytoplasmic transport deficits in transgenic flies harbouring the repeats. Freibaum et al. (2015) generated flies harbouring 58 G_4C_2 repeats in which

poly-GP and poly-GR (but not poly-GA, poly-PR or poly-PA) was detectable. These flies demonstrated both motor deficits and retinal degeneration. Using these 58 repeat flies Freibaum et al. (2015) performed a chromosomal deficiency screen spanning the entire 2nd and 3rd chromosomes. The authors identified numerous proteins involved in nucleocytoplasmic transport as enhancers and suppressors of the 58 G_4C_2 repeat induced degenerative eye phenotype. One of the strongest enhancers was the nuclear pore protein Nup50, whilst the strongest suppressor was *Ref1*, the *Drosophila* homologue of ALYREF, an RNA binding protein which acts in a complex to deliver processed mRNAs to the nuclear pore (Freibaum et al., 2015). Indeed numerous other nucleocytoplasmic transport proteins were identified as suppressors and enhancers of 58 G_4C_2 repeat degeneration including *Crm1, Nup153, Nup107, Nup106 Transportin, Gle1* and *Ran.* Further characterisation of *Nup107* morphology in *Drosophila* salivary glands revealed the protein, which in controls labels the nuclear envelope, forming a distinct nuclear boundary, appeared wrinkled and jagged and also formed inclusions near the nuclear envelope in the presence of the repeats. Morphological deficits in the nuclear envelope were further visualised by Lamin C staining, in which an abnormal frayed nuclear envelope was seen in over 40% of cells expressing the 58 G_4C_2 repeats. Additionally 58 G_4C_2 repeat cells showed a nuclear retention of RNA. The authors of this study were ambivalent as to whether it was the DPRs or G_4C_2 RNA that caused these nuclear pore deficits.

Confirmation that the arginine DPRs are indeed sufficient to disrupt nucleocytoplasmic transport came from a yeast screen using poly-PR only alternative codon constructs by Jovičić et al. (2015). By combining the results from a gain and loss of function screen Jovičić et al. (2015) found an overrepresentation of genes involved in nucleocytoplasmic transport as modifiers of poly-PR toxicity. Some of the strongest suppressors of poly-PR toxicity were an upregulation of yeast karyopherin genes. Survival of rodent cortical neurons transfected with poly-PR was more than doubled when co-transfected with KPNA3. Other notable modifiers included the yeast homologue of human RCC1, which was depleted from the nucleus in 70-80% of *C9ORF72* iPSCs. The link between the arginine DPRs and nucleocytoplasmic transport was further elucidated in a *Drosophila* screen by Boeynaems et al. (2016a).

The authors performed an RNAi screen to look for modifiers of poly-PR toxicity using *Drosophila* homologues of the genes identified from the Jovičić et al. (2015) yeast screen. Nuclear pore complex (NPC) proteins, importins, exportins, regulators of the Ran-GTP cycle and arginine methylases were all modifiers of poly-PR retinal degeneration. The strongest enhancer was the *Drosophila* homologue of the importin TNPO1, *Trn*. Boeynaems et al. (2016a) proposed that poly-PR may compete with *Trn* for *Trn* cargoes, and indeed the neuronal RNA binding protein *Elav*, a *Trn* cargo, was depleted from the nucleus and accumulated in the cytosol of *Drosophila* retinas expressing poly-PR. Furthermore, the TNPO1 cargo hnRNP A3 was aggregated in patient C9FTD brain. In addition, the arginine methyltransferase PRMT1 co-localised with both poly-GR and poly-PR in cells; and the arginine methyltransferase ASYM24 was able to detect asymmetric arginine dimethylation of poly-GR. Methylated inclusions using the ASYM24 antibody were also detected in human post-mortem tissue. It is worth noting that TNPO1 is involved in shuttling of FUS between the cytoplasm and nucleus (Dormann et al., 2012) and PRMT1 is associated with FUS (Dormann et al., 2012); however, FUS pathology is absent from the *C9ORF72* ALS/FTD brain. TNPO1 binds to PY-NLS motifs (Dormann et al., 2012) to initiate the nuclear import of proteins such hnRNPA1 (Mihevc et al., 2017). Khosravi et al. (2017) using an RFP-NLS_{PY} containing the PY-NLS of hnRNP A1 found that none of the arginine DPRs disrupts this import pathway as the nuclear localisation of the reporter was unchanged following co-expression with the reporter. It must be noted however that FUS has been shown to interact distinctly differently with TNPO1 as compared to other PY NLSs' (Niu et al., 2012).

Mechanistic insight into how poly-PR might disrupt nucleocytoplasmic transport has come from Shi et al. (2017). Shi et al. (2017) found poly-PR binds to the central channel of the nuclear pore, binding to polymeric forms of the phenylalanine:glycine (FG) repeat domain. By binding to the FG repeat domain poly-PR shifts these FG repeats toward a polymerised state. By either mutating or melting the FG domain polymers using aliphatic alcohols, poly-PR was unable to bind to nuclear pore proteins. The authors suggest the FG repeats of nuclear pore proteins exist in equilibrium between the polymerized and unpolymerized state; poly-PR binds to the FG repeats thereby stabilizing them in the polymerized state, this change in

the equilibrium of FG repeat polymerization disrupts transport through the nuclear pore, potentially by making the barrier of central channel less permeable.

1.9.3.2 Poly-GR and poly-PR associated nucleolar dysfunction

Studies investigating poly-GR and poly-PR from alternative codon ATG constructs revealed the peptides have a nucleolar localisation in cells (Kwon et al., 2014; Wen et al., 2014; Zhang et al., 2014; May et al., 2014; Tao et al., 2015; Lee et al., 2016). The nucleolus is required for ribosomal RNA (rRNA) transcription, pre-rRNA processing and ribosome subunit assembly (Lam et al., 2005). Kwon et al. (2014) provided the first evidence showing the arginine DPRs disrupt nucleolar function; observing the ability of nucleoli to synthesize mature rRNA was dramatically reduced following poly-PR expression. Jovičić et al. (2015) found in addition to proteins involved in nuclear import and export, genes involved in rRNA processing were also potent modifiers of poly-PR toxicity in yeast. Indeed, the authors discovered by restoring the function of rRNA processing machinery poly-PR toxicity can be overcome, suggesting deficits in the rRNA processing pathway are involved in arginine DPR toxicity. Additionally, in the Boeynaems et al. (2016a) *Drosophila* screen investigating modifiers of poly-PR toxicity the authors identified numerous genes linked to rRNA processing. For example, the *Drosophila* homologue nucleolin, a nucleolar RNA binding protein also identified in the Jovičić et al. (2015) yeast screen, was one the strongest modifiers of poly-PR toxicity identified. Tao et al. (2015) found the both arginine DPRs, but not any of the other RAN peptides cause the translocation of the nucleolar stress sensor protein B23 to the nucleoplasm in addition to nucleolar swelling. The maturation of 18S and 28S rRNA was also perturbed by these arginine DPRs (Tao et al., 2015).

Combined these studies suggest nucleolar stress resulting from actions of poly-GR and poly-PR may contribute to disease pathogenesis. The clinical relevance of nucleolar stress in C9ALS/FTD was recently investigated in the human brain. Mizielinska et al. (2017) found neuronal nucleoli in the C9FTD brain were significantly smaller compared to controls. However, in C9FTD brains, neurons which contained a poly-GR inclusion had a significantly increased nucleolar volume compared to cells without a poly-GR inclusion. Overexpression of poly-GR (and poly-GA but to a much lesser extent) caused nucleolar enlarged in *Drosophila* neurons.

Mizielinska et al. (2017) suggest this nucleolar enlargement may represent a disruption in the normal physiological structure of the organelle. Nucleolar dysfunction has also been associated with repeat RNA toxicity. Haeusler et al. (2014) found nucleolin preferentially binds the sense G_4C_2 RNA G-quadruplex. Nucleolin was found to be sequestered into RNA foci in the motor cortex of expansion carriers. Mizielinska et al. (2017) also observed a significant increase in nucleolar volume in C9FTD frontal cortex neurons containing sense RNA foci. This data taken together suggests a synergism between RNA and arginine DPR toxicity with regards to nucleolar dysfunction.

1.9.3.3 Poly-GR and poly-PR associated mis-splicing

As discussed in section 1.8.2 mis-splicing associated with the *C9ORF72* mutation have been attributed to the sequestration of splicing factors by the G_4C_2 RNA (Cooper-Knock et al., 2014; Haeusler et al., 2014; Cooper-Knock et al., 2015; Conlon et al., 2016). Mis-splicing events also occur as a result of arginine DPR expression. Kwon et al. (2014) cultured human astrocytes and exposed them to poly-PR for 6 hours. RNA sequencing analysis revealed changes in splicing to a number of mRNAs. Splicing errors resulting from poly-PR included exon skipping, different 5' UTRs and intronic retention. Skipping of exon 9 of the excitatory amino acid transporter 2 (EAAT2) and the inclusion of intronic sequence downstream from the splice donor site of exon 7 is observed in ALS patients (Lin et al., 1998). Cultured human astrocytes incubated with poly-PR showed the same disrupted EAAT2 splicing alterations in a poly-PR concentration dependent manner. This alteration in EAAT2 splicing was specific to poly-PR as other toxins failed elicit changes in EAAT splicing in cultured human astrocyte cells (Kwon et al., 2014).

The mechanisms of arginine DPR induced mis-splicing have been further elucidated by Yin et al. (2017). Yin et al. (2017) found when either poly-GR or poly-PR was added to nuclear extracts splicing was dramatically blocked. Splicing intermediates exon 1 and lariat-exon 2 were not detected suggesting that this inhibition of splicing occurs prior to the first catalytic step of splicing. Further poly-GR and poly-PR seem to specifically block splicing, as other stages of gene expression and RNA processing - such as primary mircoRNA processing, transcription, U6 snRNA processing, and tRNA processing were not affected.

Yin et al. (2017) found both poly-GR and poly-PR associate with the U2 small nuclear ribonucleoprotein particle complex (U2 snRNP). All known U2 snRNP snRNP components were found from proteomic analysis of poly-GR and poly-PR pulldowns. On the other hand only 3 of 8 components specific for U5 snRNP were found in the pulldowns. Hence poly-GR and poly-PR specifically associate with U2 snRNP components. Ying et al. (2017) go on to demonstrate the arginine DPRs disrupt normal U2 snRNP function. Further poly-GR inhibited the formation of the spliceosome leading to mis-splicing. Yin et al. (2017) also found the U2 snRNP associated proteins were mislocalised to the cytoplasm in *C9ORF72* iPSC motor neuron whereas it was completely nuclear in controls. On the other hand, U1 snRNP proteins were properly nuclear in both patient and control iPSCs motor neurons. In additon U2 snRNP dependent exons were found to be preferentially mis-spliced in C9 patients. Furthermore mis-splicing resulting from hnRNP H dysfunction, attributed to G_4C_2 RNA toxicity by Conlon et al. (2016) was significantly less frequent that mis-splicing resulting from U2 snRNP dysfunction in the C9ALS cerebellum and frontal cortex. This data strongly suggests DPR induced U2 snRNP dysfunction plays the crucial role in the mis-splicing observed in C9 patient brains rather than mis-splicing deficits that have been reported to result from the sequestration of RNA binding proteins into G_4C_2 foci.

1.9.3.4 Poly-GR and poly-PR associated DNA damage and oxidative stress

Lopez-Gonzalez et al. (2016) generated multiple iPSC lines differentiated into spinal cord motor neurons that produced both RNA foci and detectable levels of DPRs – including poly-GR. Analysis of these iPSC motor neurons revealed increased expression of γH2AX, a marker of DNA double-strand breaks (DSBs). DNA damage activates the p53 pathway, and indeed p53 levels were also increased in C9 patient motor neuron iPSCs compared to controls. Poly-GR expression activated the p53 pathway, increased levels of γH2AX and increased the number of DNA strand breaks when transfected into control iPSC motor neurons. Further the toxicity of poly-GR in *Drosophila* was attenuated by reducing p53 levels. Next the authors assessed how poly-GR might be causing DNA damage, the production of reactive oxygen species (ROS) was significantly increased from 8 weeks onwards in C9 iPSC

motor neurons compared to controls. When poly-GR was expressed in HEK293 cells and interactome analysis performed – many of the top interactors were mitochondrial ribosomal proteins necessary for the translation of 13 subunits of the mitochondrial complexes. Mitochondrial membrane potential was found to be elevated in C9 iPSC motor neurons and control iPSCs transfected with poly-GR. This data suggests poly-GR perturbs mitochondrial function by binding to mitochondrial ribosomal proteins. DNA damage has been reported in the spinal cord neurons of *C9ORF72* ALS patients (Farg et al., 2017). Farg et al. (2017) found multiple markers of DNA damage such as γH2AX and p-ATM are significantly upregulated in *C9ORF72* patients compared to controls. This research taken as a whole indicates neuronal poly-GR disrupts mitochondrial function and induces DNA damage potentially via elevated oxidative stress. Further work is needed however to determine what other molecular pathways may contribute to DNA damage in C9ALS/FTD.

1.9.3.5 Poly-GR and poly-PR associated ER Stress and perturbed cell signalling pathways

A recent CRISPR-Cas9 screen in human cells and primary neurons has revealed ER stress as a potent modifier of arginine DPR toxicity (Haney et al., 2017). ER-resident proteins such as TMX2 and CANX were strong modifiers of poly-GR and poly-PR toxicity; virtually all of these proteins were members of the endoplasmic reticulum membrane protein complex (EMC). Endosomal trafficking genes such as Rab7a were also potent poly-PR interactors. The large number of ER associated genetic modifiers indicates poly-GR and poly-PR both induce an ER stress response, akin to what has been reported for poly-GA toxicity (Zhang et al., 2014; May et al., 2014). To confirm this Haney et al. (2017) performed an RNA-seq on primary neurons transduced with poly-PR. The most significantly enriched Gene Ontology analysis for upregulated genes was the ‘apoptotic signalling pathway in response to endoplasmic reticulum stress’. ER stress genes such as ATF4 were upregulated in poly-PR cells; further poly-PR toxicity was mitigated by treating cells with an ER stress inhibitor that prevents Atf4 induction. A PERK-Atf4 mediated ER stress response has been reported in C9ALS/FTD patient brains by Zhang et al. (2014).

1.9.3.6 Poly-GR and poly-PR associated disruptions in global protein translation

Kanekura et al. (2016) synthesized a C-terminally tagged 20 repeat constructs for poly-PR and performed an immunoprecipitation analysis followed by a liquid chromatography-tandem mass spectrometry (LC-MS/MS) to identify poly-PR interacting proteins. The most abundant interactors were ribosomal proteins, proteins involved in splicing and translation initiation and elongation factors indicating poly-PR may influence protein translation (Kanekura et al., 2016). Protein translation was inhibited following poly-PR and poly-GR expression, but not poly-GA. However, poly-PR did not alter the phosphorylation levels of eIFs, the phosphorylation status of which regulates protein translation. Poly-PR was found to potentially form complexes/aggregates with RNA, further RNA induced aggregate formation of both poly-PR and poly-GR. Poly-PR/RNA complexes were recognised as misfolded proteins *in vitro* by HSP70; together these findings suggest the arginine DPRs may prevents the access of translation factors to mRNA by forming complexes with said mRNA; thereby blocking protein synthesis.

1.9.3.7 Poly-GR and poly-PR disruption of the formation of membraneless organelles and their liquid-liquid phase separation dynamics

The binding partners of poly-GR and poly-PR were revealed initially by proteomic analysis by Lee et al. (2016) and Lin et al. (2016). These studies revealed a strong interaction between poly-GR and poly-PR with proteins that contain a low complexity sequence domain (LCD). LCDs are amino acid sequences between 75-300 amino acids in length with a high evolutionary conservation (Freibaum and Taylor, 2017). LCDs are present in one-third of the human proteome, and are predicted to be unstructured and consist of a high number of glycine and serine residues interspersed with aromatic charged residues (Freibaum and Taylor, 2017). LCDs are involved in multivalent, low affinity interactions; hence mediating interactions that can be rapidly rearranged; including both protein-protein and protein-nucleic acid interactions (Aguzzi and Altmeyer, 2016); these interactions are important for several cellular functions – including the formation of dynamic heterogeneous assembles; such as membraneless organelles (Aguzzi and Altmeyer, 2016; Freibaum and

Taylor, 2017). Membraneless organelles are multicomponent, viscous liquid-like structures that typically contain both RNA and protein molecules and are commonly referred to as RNP bodies or granules (Brangwynne et al., 2015). Examples of such membraneless organelles include the nucleolus, nuclear pore complex, stress granules nuclear speckles and Cajal bodies (Brangwynne et al., 2015; Freibaum and Taylor, 2017).

Membraneless organelle formation occurs through spontaneous liquid-liquid phase separations (LLPS); a process in which protein-laden RNAs separate themselves from the surrounding aqueous nucleoplasm or cytoplasm (Brangwynne et al., 2015; Taylor et al., 2016). The formation of membraneless organelles has been described as dynamic liquid demixing process in which cells actively generate phase boundaries to confine functional entities for temporary period (Aguzzi and Altmeyer, 2016). Once a phase transition has occurred it is thought the proteins within the liquid compartment find themselves in different solvent environment compared to their surroundings, which may promote specific biochemical interactions (Aguzzi and Altmeyer, 2016). Most of the proteins that drive intracellular phase transitions, and the formation of membraneless organelles, show strong conformational heterogeneity and are referred to together as intrinsically disordered proteins (IDPs) (Brangwynne et al., 2015; Aguzzi and Altmeyer, 2016). IDPs do not have well defined protein fold like other peptides and are highly flexible; this structural plasticity allows IDPs to dynamically adopt different confirmations and undergo a multitude of promiscuous multivalent interactions (Aguzzi and Altmeyer., 2016).

Many IPDs contain LCDs, as noted previously LCDs form multivalent interactions and generate energetically favourable higher-order protein assemblies; such factors dictate the behaviour of IPDs within complex liquids like the intracellular matrix, thereby facilitating phase transitions (Brangwynne et al., 2015; Aguzzi and Altmeyer, 2016). RNA binding proteins such as FUS, TDP-43 and hnRNPs contain prion-like domains (PrLD), a type of LCD that is rich in uncharged polar amino acids and has the propensity to self-assemble and aggregate (Taylor et al., 2016; Aguzzi and Altmeyer, 2016; Boeynaems et al., 2017). FUS, recombinant TDP-43 and hnRNPs have been shown to undergo liquid-liquid phase separations (Boeynaems et al., 2017) in which they transition from a single mixed aqueous solution into two distinct phases; one of which is a concentrated liquid-like droplet (Taylor et al., 2016;

Boeynaems et al., 2017). Taylor et al. (2016) note during LLPS RNA binding proteins which are prone to fibrillize are placed at a high concentration in close proximity to one another increasing the chance of these proteins forming amyloid-like fibrils in the cytosol. Indeed Boeynaems et al. (2017) note liquid like droplets mature over time to more fibrillary states. ALS causing mutations in proteins such as TDP-43, FUS and hnRNPA1 are often found in the PrLD of these proteins (Patel et al., 2015; Murakami et al., 2015; Molliex et al., 2015; Lin et al., 2015). Taylor et al. (2016) notes such mutations accelerate the transition of liquid like droplets to pathological amyloid-like fibrils. Indeed, the formation of pathological aggregates of RNA binding proteins is proposed to be a liquid-to-solid phase transition (Molliex et al., 2015; Patel et al., 2015).

The molecular signals that trigger phase transitions and how protein assembly into membraneless compartments is regulated is not fully known but will be important for understanding how perturbed phase transitions transform liquid compartments into the fibrous aggregates seen in neurological diseases (Aguzzi and Altmeyer, 2016). Poly-GR and poly-PR were shown by Boeynaems et al. (2017) to undergo LLPS *in vitro*, LLPS correlated with arginine content of peptides indicating the arginine residues within these DPRs module their LLPS. Further characterisation of poly-PR strongly indicated poly-PR droplets behave as a liquid, with poly-PR droplets existing in a disordered state. Both poly-GR and poly-PR associate with LCDs; Lin et al. (2016) demonstrated poly-PR directly binds to LCDs and the LCD is both necessary and sufficient for poly-PR binding and this interaction is polymer-dependent; indicating the polymeric structure of the arginine DPRs is the toxic species (Freibaum and Taylor, 2017). Mass spectrometry by Boeynaems et al. (2017) demonstrated poly-PR interacted with proteins enriched for RNA-binding domains, arginine-rich motifs and protein disorder; many of which were components of membraneless liquid organelles.

Important examples of IDPs shown to interact with poly-GR and poly-PR are RNA-binding proteins associated with or known to cause ALS including FUS, Ataxin-2, Matrin-3, hnRNPA2B1 and importantly TDP-43 (Lee et al., 2016; Lin et al., 2016). Based on the finding that poly-GR and poly-PR interactors are significantly enriched in RNA binding proteins with LCD domains (Lee et al., 2016; Lin et al., 2016) involved in the formation of membraneless organelles, Lee et al. (2016) performed a

Drosophila RNA interference (RNAi) screen in poly-GR expressing flies. 84.9% of the poly-GR/PR interactors found in the proteomics analysis by Lee et al. (2016) were genetic modifiers of poly-GR toxicity. Knockdown of these LCD interacting proteins revealed 80 suppressors, 35 of which were strong suppressors of poly-GR toxicity. 27 enhancers were found, 21 of which were strong enhancers of poly-GR toxicity. Genetic modifiers from the poly-GR screen were also found to modify the rough eye phenotype of $(G_4C_2)_{58}$-expressing flies, shown prior to produce poly-GR (Freibaum et al., 2015). Previous studies had shown both poly-GR and poly-PR associate with membraneless organelles as the peptides accumulate in the nucleoli of cells (Kwon et al., 2014; Wen et al., 2014; Zhang et al., 2014; May et al., 2014; Tao et al., 2015). Consistent with these previous studies, Lee et al. (2016) also observed this; with poly-PR being exclusively nuclear and poly-GR being both nuclear and cytoplasmic. Lee et al. (2016) demonstrated poly-GR and poly-PR are recruited to the liquid-like granular component (GC) of the nucleolus; although poly-GR shows a more stable interaction with the nucleolus potentially due to its interaction with dense fibrillar components of the nucleolus. The liquid-like GC component of the nucleolus is organised by nucleophosmin (NPM1) (Lee et al., 2016). NPM1 contains 3 LCDs and undergoes LLPS to form liquid droplets (Mitrea et al., 2016; Lee et al., 2016). Poly-GR and poly-PR both directly interacted with NPM1 and were sufficient to induce LLPS of NPM1. Further poly-GR and poly-PR modulated the ability of NMP1 to undergo LLPS with one of its native nucleolar binding partners SURF6. Poly-GR and poly-PR in increasing concentrations reduced the concentration of SURF6 required to induce LLPS; and when the concentration of the arginine DPRs was greater than NPM1, SURF6 could no longer induce LLPS indicating they outcompete SURF6 for NPM1 binding. A consequence of this nucleolar interaction was a disruption in nucleolar dynamics, nucleolar morphology and nucleolar function. The motility of both NPM1 and NCL were significantly reduced, the nucleolus of cells expressing either poly-GR or poly-PR was perturbed and rRNA synthesis was reduced.

Lee et al. (2016) then focused on a cytoplasmic membraneless organelle – stress granules (SGs). SGs are cytoplasmic assemblies formed of RNA and protein during cellular stress (Protter and Parker, 2016). RNA binding proteins are bound by increased levels of free mRNA following translational arrest which nucleates SG

formation (Jain et al., 2016). SG formation requires LLPS (Lin et al., 2015; Murakami et al., 2015). Lee et al. (2016) found both poly-GR and poly-PR interact with numerous SG components including G3BP1 and G3BP2 (G3BP) and its binding partner Caprin1 – both proteins promote SG assembly. Poly-GR (but not poly-PR) co-localised with SG markers, although the expression of either peptide led to spontaneous SG assembly in cells. Furthermore, these SGs were poorly dynamic, rarely disassembled and their numbers increased over time. Further G3BP when in SGs is normally in a rapid dynamic equilibrium with the cytoplasm; however, poly-GR and poly-PR impaired the exchange of G3BP between SGs and the cytoplasm. These arginine DPR induced SGs were associated with increased risk of cell death. The translation of mRNA, stalled upon SG formation, was significantly reduced in poly-GR and poly-PR expressing cells demonstrating further SG dysfunction. Interestingly TDP-43 was also recruited to poly-GR and poly-PR induced SGs (Lee et al., 2016). The interaction between poly-PR and SGs has been further elucidated by Boeynaems et al. (2017), finding cytoplasmic poly-PR granules are positive for SG markers; further poly-PR induced SG assembly was both dose and length dependent. The phosphorylation of eIF2α is required for SG assembly (McInerney et al., 2005), using a mutant MEF cell line which has a non-phosphorylatable form of eIF2α, Boeynaems et al. (2017) found poly-PR requires the integrated stress granule response for SG induction. Additionally, G3BP was also required for poly-PR induced SG formation. Further ALS associated proteins with LCDs, ataxin-2 and TDP-43, were both enriched in poly-PR induced SGs. Tao et al. (2015) observed poly-GR and poly-PR can also prevent SG assembly in response to arsenite induced stress.

Other membraneless organelles are also perturbed by poly-GR and poly-PR. Lee et al. (2016) found the assembly of Cajal bodies and the dynamics of nuclear speckles was perturbed by these arginine-rich DPRs. Poly-GR and poly-PR also influence the biophysical properties of LCD containing RNA binding proteins. Lee et al. (2016) found poly-GR and poly-PR were recruited to hnRNP and TIA-1 liquid droplets and decreased the critical concentration required for hnRNP and TIA-1 to undergo LLPS. Further, the dynamic exchange of hnRNPA1 and TIA-1 between the dense droplet phase and the surrounding mono-disperse phase was perturbed by poly-GR and poly-PR leading Lee et al. (2016) to suggest these DPRs enhance the multivalent

interactions that comprise the liquid phase of RNA binding proteins (figure 1.19). Additionally, Boeynaems et al. (2017) found poly-PR, but not poly-GR interfered with the spontaneous fusion dynamics of FUS liquid droplets, and FUS LLPS was enhanced in the presence of poly-PR.

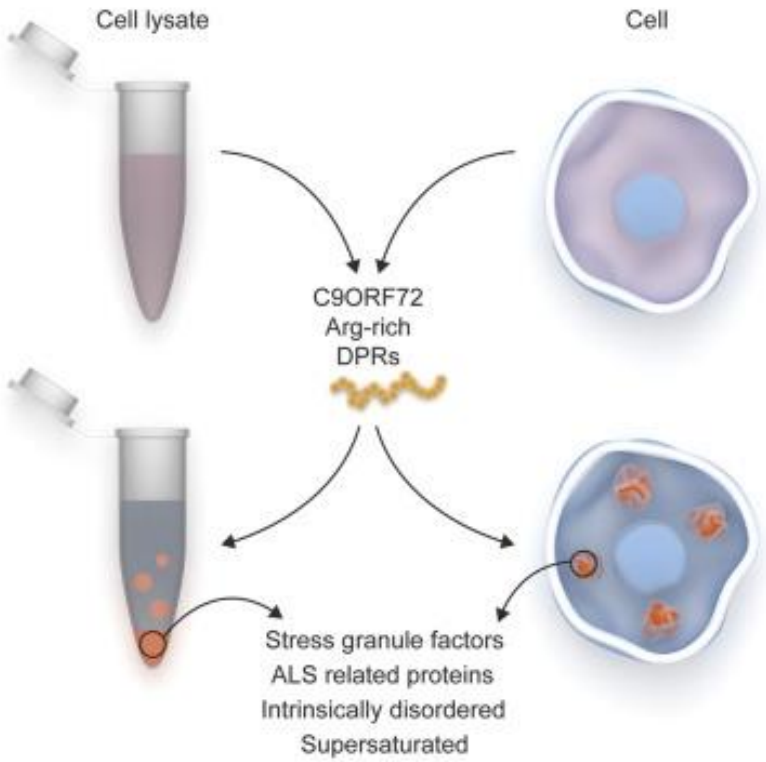

Figure 1.19 | Arginine-rich DPR phase separation. The arginine rich DPRs, poly-GR and poly-PR, undergo liquid-liquid phase separation in addition to also inducing phase separation of numerous proteins harbouring low complexity domains, such as intrinsically disordered ALS related proteins involved in RNA metabolism and stress granule formation. Poly-GR and poly-PR when expressed in cells causes stress granules to spontaneously form. From Boeynaems et al., (2017).

In summary experimental data demonstrates these DPRs promote phase separation of PrLD proteins into liquid droplets. The interactions mediated by LCDs that permits LLPS to occur is unknown; although it is thought arginine rich motifs are important potentially due to the ability of this residue to participate in charge-charge and Pi-cation interactions within the multivalent network (Lee et al., 2016; Nott et al., 2015; Pak et al., 2015). Indeed Boeynaems et al. (2017) demonstrated the importance of electrostatic forces via arginines in LLPS; finding phase transitions correlated with the arginine content of peptides hence supporting the function of arginine motifs in mediating LLPS. Hence the arginine DPRs likely form strong multivalent interactions with proteins harbouring LCDs (Lee et al., 2016); indeed, in the case of NPM1, poly-GR and poly-PR outcompete the normal binding partner of NPM1, SURF6 for binding to NPM1 (Lee et al., 2016). Such strong multivalent interactions make LLPS easier to occur essentially causing excessive droplet formation. The dynamics and maturation of the liquid droplets formed by PrLD containing proteins is also disrupted by the binding of the arginine DPRs. The arginine DPRs appear to make liquid droplets more viscous by enhancing these multivalent interactions occurring within the liquid phase of PrLD droplets (Lee et al., 2016). This increased viscosity will likely enhance the maturation of liquid droplets to a less dynamic and more rigid state, as has been shown for the liquid droplets formed by FUS and hnRNPs (Molliex et al., 2015; Patel et al., 2015). In support of this Boeynaems et al. (2017) found poly-PR accelerates the liquid-to-solid maturation of PrLD droplets. Further the strong interaction between poly-GR and poly-PR with the LCD disrupts the dynamic exchange of proteins such as FUS and hnRNPA1 between the dense droplet phase and the surrounding mono-disperse phase as previously noted (Lee et al., 2016). The result of these perturbed and poorly dynamic phase transitions is that the liquid compartments of cellular structures that form via phase separations will be disturbed. The most obvious consequence is the internal functional balance of the liquid components of membraneless organelles will be disrupted (Boeynaems et al., 2017).

1.9.3.8 Summary of poly-GR and poly-PR toxicity

In summary the arginine DPRs are highly toxic to neurons, particularly in *Drosophila* models (Mizielinska et al., 2014; Yang et al., 2015; Freibaum et al., 2015; Boeynaems et al., 2016a). The altered assembly, dynamics, and function of

membraneless organelles that result from disturbed phase transitions fully account for the widespread cellular abnormalities observed in C9ALS/FTD and can explain several of the other mechanisms of poly-GR and poly-PR toxicity. The splicing alterations induced by these arginine-rich DPRs observed by Kwon et al. (2014) and Yin et al. (2017) is potentially attributable to the alteration of liquid-like properties in nuclear speckles by poly-GR and poly-PR (Lee et al., 2016; Freibaum and Taylor, 2017). Nuclear speckles are enriched in pre-mRNA splicing factors (Spector and Lamond, 2011); hence a disruption in their dynamics is likely to affect splicing (Lee et al., 2016). Indeed Yin et al. (2017) note U2 snRNP proteins were identified as poly-GR and poly-PR interactors by Lee et al. (2016) and that many of these U2 snRNP proteins contain LCDs. Yin et al. (2017) suggest the interaction between these arginine DPRs and LCD containing U2 snRNP proteins may explain why they observed a loss of U2 snRNP proteins from nuclear speckles and an accumulation in the cytosol. The nucleolar dysfunction observed following arginine DPR expression (Kwon et al., 2014; Tao et al., 2015; Mizielinska et al., 2017) is also caused by the disruption of the liquid properties of the nucleolus (Lee et al., 2016). LLPS is also important for the transport of molecules across the NPC (Schmidt and Görlich, 2016). As discussed previously both poly-GR and poly-PR disrupts nucleocytoplasmic transport (Jovičić et al., 2015; Boeynaems et al., 2016; Lee et al., 2016); further poly-PR was shown by Shi et al. (2017) to bind to and stabilize polymeric FG repeats found in the core of the nuclear pore changing the properties of the central channel of nucleoporins (NUPs). Hence poly-GR and poly-PR induced LLPS deficits may also perturb nucleocytoplasmic transport. Indeed Shi et al. (2017) note the sequences found in the FG domains of nucleoporins (repetitions of the tripeptide sequence glycine/serine-tyrosine-glycine/serine) are like the LC domains found in RNA binding proteins. Low-complexity FG domains have been reported to phase separate via multivalent cohesion to form a sieve like selective hydrogel barrier (Aguzzi and Altmeyer, 2016). The arginine DPRs appear to act as molecular seeds initiating phase transitions by binding to the LCD of intrinsically disordered proteins causing liquid demixing; furthermore, due to the strength of the interaction between these DPRs and LCDs, the liquid droplets formed become more viscous thereby perturbing their normal biological functions (figure 1.20).

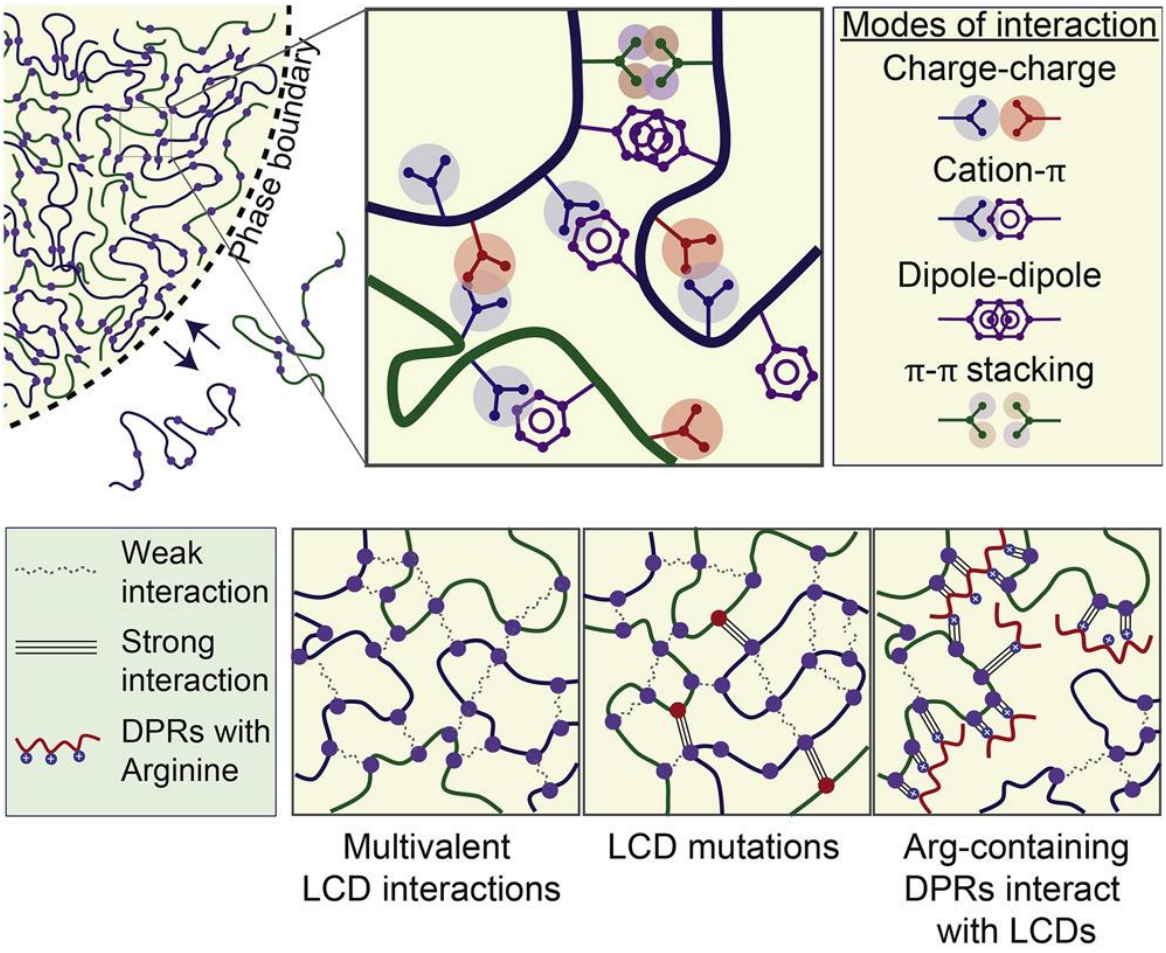

Figure 1.20 | LCD mutations and the arginine DPRs disrupt LCD interactions. Multivalent interactions drive phase separations involved in the formation of membraneless organelles. Mutations in the low complexity domain (LCD) of proteins and the arginine DPRs (poly-GR and poly-PR) derived from the G_4C_2 expansion alter the strength of LCD interactions driving phase separations and making the liquid droplets resulting from these phase separations more viscous due to the strength of the interactions formed. From Lee et al. (2016).

1.9.4 Poly-GP toxicity

Poly-GP is the second most abundant DPR after poly-GA in the human C9 ALS/FTD brain (Schludi et al., 2015). Poly-GP is uncharged and has a compact flexible coil structure (Freibaum and Taylor, 2017). Poly-GP is unique amongst the DPRs as it is produced from both the sense and anti-sense strand, with the anti-sense poly-GP species surprisingly appearing the more abundant species (Zu et al., 2013). Flores et al. (2016) demonstrated the poly-GP peptide is resistant to aggregation and does not appear to associate with markers of amyloid aggregates consistent with cellular models using poly-GP only constructs (Zhang et al., 2014; May et al., 2014; Wen et

al., 2014; Tao et al., 2015; Yamakawa et al., 2015). Mass spectrometry analysis has revealed poly-GP interacts with few cellular proteins (Lee et al., 2016).

Evidence for poly-GP toxicity is mixed; with mild (Zu et al., 2013; Yamakawa et al., 2015) to no toxicity being reported in several studies (Zhang et al., 2014; May et al., 2014; Wen et al., 2014; Tao et al., 2015). Yamakawa et al, (2015) found poly-GP impaired UPS degradation. Using the UPS reporter Ub-G76V-GFP, which is normally rapidly degraded by the UPS Yamakawa et al. (2015), observed Ub-G76V-GFP levels were higher in poly-GP transfected cells. Poly-GP also enhanced cell death resulting from exposure to the proteasome inhibitor MG132. Similar results were seen for poly-GA and poly-GR. Zu et al. (2013) using a G_4C_2 construct that overexpressed poly-GP (via an insertion of an ATG in the poly-GP frame) found increased cell death. In the human brain poly-GP levels associate with neuropathological subgroup, poly-GP levels are significantly higher in the frontal cortex of ALS/FTD patients compared to FTD frontal cortex poly-GP levels (Gendron et al., 2015). Additionally, poly-GP levels are significantly lower in ALS patient cerebellum compared to both ALS/FTD and FTD cerebellum. Interestingly Gendron et al. (2015) did observe cerebellar poly-GP associates with cognitive impairment in ALS and ALS/FTD patients. However, poly-GP levels in frontal cortex, motor cortex and the hippocampus did not correlate with cognitive impairment. Furthermore poly-GP did not associate with age of disease onset or survival after onset (adjusted for age of onset and disease subgroup) in any of the previously mentioned brain regions.

1.9.4.1 Poly-GP as a pharmacodynamic biomarker for C9ALS/FTD

Whilst poly-GP toxicity is unlikely to be the primary pathogenic species in C9ALS/FTD it does appear to represent a useful biomarker for C9ALS/FTD (Su et al., 2014; Gendron et al., 2017; Lehmer et al., 2017). Su et al. (2014) developed a poly-GP immunoassay to measure soluble levels of poly-GP in cerebrospinal fluid (CSF); poly-GP was only detectable in C9 ALS CSF, but not SALS or control CSF. The confirmation of poly-GP as a pharmacodynamic biomarker for C9ALS/FTD has come recently from Gendron et al. (2017) and Lehmer et al. (2017). Gendron et al. (2017) replicated the previous findings from Su et al. (2017) that poly-GP is detectable in C9ALS CSF, however went further in showing that the peptide is

detectable in CSF of both symptomatic and asymptomatic C9 ALS and C9ALS/FTD expansion carriers. Further characterisation of poly-GP CSF levels in asymptomatic expansion carriers by Gendon et al. (2017) revealed CSF poly-GP is stable over time. However, poly-GP CSF levels were not a prognostic marker for either C9ALS or C9ALS/FTD; CSF levels of poly-GP did not correlate with age at disease onset, survival after disease onset, onset site or disease group. No correlation was found between poly-GP CSF levels and behavioural impairment. CSF poly-GP was associated with cognitive impairment however the significance of this finding was lost when the authors adjusted for age at CSF collection, gender, and years of education (Gendron et al. 2017). Poly-GP was also detectable in the lymphoblastoid cell lines from C9 expansion carriers and in the media of these cells indicating poly-GP is actively secreted from cells. Gendron et al. (2017) also found poly-GP was a useful marker to test for the effectiveness of ASOs targeting the G_4C_2 RNA. A lymphoblastoid cell line from a C9ALS patient and a line from an asymptomatic C9 carrier treated with the G_4C_2 ASO reduced poly-GP levels. This G_4C_2 ASO reduced both intracellular and extracellular poly-GP in C9 iPSC neurons. Furthermore, these decreases paralleled reductions in G_4C_2 RNA; these findings together indicate extracellular poly-GP in the CSF can also be used to determine the effectiveness of potential treatments. The detection of poly-GP in CSF was confirmed in a second study by Lehmer et al. (2017). Hence these studies strongly support a role for poly-GP as a diagnostic marker for C9ALS/FTD; further poly-GP CSF levels may also provide a means to assess the effectiveness of therapies targeting the G_4C_2 RNA in expansion carriers.

1.9.5 Poly-PA toxicity

Poly-PA is the final DPR translated from the repeat expansion. Poly-PA is translated from the anti-sense C_4G_2 strand. Poly-PA inclusions are exceptionally rare (Schludi et al., 2015; Mackenzie et al., 2015), with only very rare inclusions being observed in the cortex or motor neurons of patients (Mackenzie et al., 2015); although poly-PA aggregates have been reported in one study to be more prevalent in motor neurons than poly-GA inclusions (Cooper-Knock et al., 2015). Poly-PA has a compact flexible coil structure and is uncharged, and consistent with its structure the peptide is seemingly inert (Freibaum and Taylor, 2016). No study utilizing alternative codons

has reported any evidence of poly-PA toxicity or that it forms inclusions (Freibaum and Taylor, 2016) with the protein being diffusely cytoplasmic in cellular models (Zhang et al., 2014; May et al., 2014; Wen et al., 2014; Yamakawa et al., 2015). Despite no evidence for poly-PA toxicity a recent study has suggested poly-PA may play a protective role by ameliorating poly-GA toxicity. Lee et al, (2017) found poly-PA is sequestered by poly-GA and that co-expression of poly-PA with poly-GA reduces poly-GA toxicity by reducing its aggregation *in vitro* and *in vivo* in chick embryos.

1.9.6 Evidence against DPR toxicity

Despite both *in vitro* and *in vivo* model systems convincingly demonstrating the potent toxicity of the DPRs, particularly poly-GR and poly-PR (Gitler and Tsuiji, 2016; Moens et al., 2017) studies of DPR distribution in the human brain are contradictory to this. Gitler and Tsuiji, (2016) note if DPRs are the driving factor behind pathogenesis then it would be predicted they would accumulate in the most vulnerable regions of the brain affected in C9ALS/FTD; and that DPR abundance should correlate with disease severity. This is however not the case; DPR pathology does not correlate with neurodegeneration or clinical severity (Davidson et al., 2014; Mackenzie et al., 2014; Davidson et al., 2016; Mackenzie et al., 2015; Schludi et al., 2015).

Mackenzie et al, (2015) performed a detailed quantitative analysis and clinic-pathological correlation of DPR proteins in C9ALS/FTD carriers. Other than poly-GA where a moderate positive association was found between the amount of poly-GA DNs in the upper cortical layers of the frontal cortex and local degeneration; no other significant correlation was found between DPR pathology and degeneration in either the frontal cortex or LMNs. Indeed, the DPRs are exceptionally rare in spinal cord motor neurons (Gomez-Deza et al., 2015; Mackenzie et al., 2015; Davidson et al., 2016) with the antisense sense DPRs poly-PA and the highly toxic poly-PR being exceptionally rare throughout the brain (Schludi et al., 2015; Mackenzie et al., 2015). Furthermore Mackenzie et al. (2015) observed no association for disease duration for any DPR protein. A moderate negative association between poly-GA pathology in the frontal cortex and age of onset was found for overall cohort; however, the authors' note this finding should be treated with caution as the correlation was driven

by two C9ALS/FTD cases with an extremely high poly-GA burden. In fact, exclusion of these two cases removed the significance of this association. Further no correlation was seen for poly-GA or any of the DPRs for age of onset for either pure FTD or pure ALS. In addition, no correlation was seen between any of the other DPRs and age of onset and no DPR protein pathology correlated with disease duration. A healthy case with a short 30 repeat unit expansion was reported by Gami et al. (2015); DPR pathology (in addition to RNA foci) was seen - suggesting the presence of DPRs is not sufficient for disease; although the abundance of DPR pathology was not reported. Furthermore, DPR toxicity does correlate with length of the dipeptide in model systems (Mizielinska et al., 2014).

A problem for the suggestion DPRs drive toxicity in C9ALS/FTD is the lack of DPR inclusions in vulnerable neuronal populations; especially the arginine DPRs, in spinal cord motor neurons (Gomez-Deza et al., 2015). Wen et al. (2014) reported the presence of poly-PR nuclear inclusions that co-localised with the nucleolus and extranuclear poly-PR in C9ALS and C9ALS/FTD spinal cords; although bizarrely the antibody used also detected poly-PR aggregates (although less frequently) in the spinal cords of non C9ALS patients and healthy controls. Cooper-Knock et al. (2015) were able to detect cytoplasmic poly-PR inclusions in motor neurons which were more numerous than poly-GA or poly-GR inclusions. Others have failed to reproduce this finding using the same poly-PR antibodies. Mackenzie et al. (2015) saw extremely rare poly-PR inclusions in the spinal cord observing fine or coarse granular nuclear labelling with larger regions of immunopositive staining seen in both C9 and controls, leading to the conclusion that the staining was non-specific. Davidson et al. (2016) reported a nuclear staining pattern in anterior horn cells similar to Mackenzie et al. (2015) with the poly-PR antibody used by Cooper-Knock et al. (2015), further Davidson et al. (2016) did not detect any poly-PR cytoplasmic or nuclear inclusions. Additionally the same staining pattern was observed in controls. Schludi et al. (2015) also reported no poly-PR spinal cord aggregates using their own antibody; Gomez-Deza et al. (2015) observed poly-PR inclusions in spinal cord motor neurons but they were exceptionally rare using the poly-PR antibody generated by Gendron et al. (2013); only one poly-PR spinal cord inclusion was seen per section was seen in 3 of the 10 C9 cases studied. Differences in antibody specificity and cell types stained likely play a role, however Davidson et al. (2016)

note Wen et al. (2014) did not specify the spinal cord cell type in which they observed poly-PR nuclear inclusions or the case numbers they investigated. Further Schludi et al. (2015) found intranuclear poly-GR and poly-PR inclusions observed in cellular models produced from alternative codon constructs do not correspond to what is seen in the human brain. Arginine DPR intranuclear inclusions derived from protein only constructs show a pronounced nuclear and nucleolar localization whereas those in the human brain are mainly cytoplasmic, and when nuclear are paranucleolar; attached next to the nucleolus - but not co-localising with it. Furthermore, overexpressed poly-GR and poly-PR (and poly-GP) are not p62 positive, whereas the majority of inclusions in the human brain are p62 positive (Schludi et al., 2015).

It has been argued that the arginine DPRs are so difficult to detect in post-mortem analysis because they are so highly toxic (particularly poly-PR), thereby the affected neurons are quickly lost and not present by the time of patient death (Gitler and Tsuiji, 2016). Further, the potently toxic poly-PR may not need to accumulate at such high levels to cause neuronal death. Gitler and Tsuiji, (2016) note the DPRs may exist in multiple conformations or strains like other disease related proteins such as tau, α–syn and Aβ, hence current DPR antibodies may not be able to detect these DPR species, thereby providing a misleading picture about DPR pathology. Differences in DPR expression pattern combined with differential aggregation properties and toxicities may influence the prevalence and distribution of DPR inclusions seen in the post-mortem brain. Edbauer and Haass, (2016) note in Alzheimer's disease (AD), the accumulation of the amyloid β-peptide (Aβ) pathology occurs decades before dementia onset in the neocortex, and then spreads centripetally toward the brain stem. Edbauer and Haass, (2016) argue by the time AD is diagnosed the Aβ pathology has plateaued and no longer correlates well with the progression of AD - as is the case with the DPRs in C9ALS/FTD. Given the poor correlation between DPR inclusion bodies and areas of neurodegeneration Mackenzie et al. (2015) investigated the soluble forms of DPR proteins and whether there is any difference in C9 soluble species between different C9 clinical manifestations. However, the authors observed no specific differences in the banding patterns of the soluble protein fractions using DPR antibodies between C9 cases, controls and C9 cases with different clinical presentations. Mackenzie et al.

(2015) interpreted this as a lack of evidence for the presence of potentially soluble toxic DPR species.

Of note, not all post-mortem analysis argues against DPR protein toxicity. Proudfoot et al. (2014) reported the case of a 26 year old expansion carrier with a severe developmental disability who died of unrelated causes and had abundant poly-GA pathology but no TDP-43 inclusions. Gendron et al. (2015) found cerebellar poly-GP levels were significantly lower in ALS patients compared to FTD or ALS/FTD; further the cerebellar levels of poly-GP correlated with cognitive score. A similar finding was observed for poly-GA levels, but this did not associate with cognitive score. However, both cerebellar poly-GP and poly-GA levels correlated with V3 mRNA expression, repeat size, disease onset and survival after onset (Gendron et al 2015). Hence cerebellar levels of the DPRs correlate with neuropathological and clinical phenotypes. Indeed, one argument that has been used to doubt the contribution of the DPRs to C9ALS/FTD pathogenesis is their high burden in regions of the brain not typically associated with ALS/FTD, particularly the cerebellum. However there have been reports of cerebellar atrophy in C9FTD cases (Bocchetta et al., 2016); further longitudinal neuroimaging study of expansion carriers revealed preferential volume loss in the thalamus and cerebellum (Mahoney et al., 2012). Cerebellar ataxia has also been reported to be associated with expansions in *C9ORF72* (Lindquist et al., 2013) and longer repeats in the cerebellum are associated with poorer survival in expansion carriers (van Blitterswijk et al., 2013). Furthermore post-mortem analyses on patients with the C9ORF72 expansion who have died early of the disease show abundant DPR pathology only; further indicating DPR pathology precedes that of TDP-43 and may initiate the disease cascade (Proudfoot et al 2014; Baborie et al 2015; Vatsavayai et al 2016; Davidson et al 2016).

Vatsavayai et al., (2016) reported on two C9 cases; the first a C9FTD case showing focal degeneration in brain regions lacking TDP-43 pathology but harbouring foci and DPRs; hence neurodegeneration can occur in the absence of TDP-43 aggregates; which normally correlates with areas of neuronal loss in C9ALS/FTD (Mackenzie et al., 2014). There was still a poor correlation between individual DPRs and RNA foci with regional patterns of neurodegeneration, suggesting that alone the specific pathological features of repeat expansion are not sufficient to cause degeneration; however, a combination of the different features in a neuron may be toxic. To

address this potential hypothesis Vatsavayai et al. (2016) looked for RNA foci and p62 positive inclusions finding neurons do indeed harbour more than one of these pathological features to different extents in multiple brain regions. This was particularly striking in the medial pulvinar thalamus (mPULV) where 27% of neurons contained sense foci, anti-sense foci and a DPR inclusion. Importantly focal neurodegeneration was seen in the mPULV. Case 2 provided an exciting opportunity to study the presymptomaic and symptomatic stages of C9FTD from the same patient. Case two was an atypical C9ALS/FTD case that had undergone a temporal lobe resection for epilepsy 5 years prior to the onset of FTD symptoms. Imaging revealed presymptomaic volume reductions suggesting early degeneration resulting from the repeat expansion. Pathological analysis of the presymptomatic tissue showed both foci and DPR pathology whereas TDP-43 pathology was absent. However, at autopsy TDP-43 pathology was abundant; interestingly the most degenerative neurons were those lacking nuclear TDP-43 staining but did not harbour a pathological inclusion. Interestingly Scaber and Talbot, (2016) note this is the fourth publication to report minimal TDP-43 pathology in C9FTD (Gijselinck et al., 2012; Proudfoot et al., 2014; Baborie et al., 2015; Vatsavayai et al., 2016) while no such cases have been reported for C9ALS. Further they note the DPRs are incredibly rare in the spinal cord; whereas they are much more common in cortical regions; hence there may be differences in DPR contribution to different clinical manifestations of the repeat expansion.

1.9.7 Summary of DPR toxicity

In summary the data from protein only DPR models *in vivo* and *in vitro* provide strong support for the toxicity of arginine DPRs, with poly-GA being identified as toxic in some model systems (Moens et al., 2017) (figure 1.21). Mechanisms of poly-GR and poly-PR toxicity have been recently elucidated (Lee et al., 2016; Lin et al., 2016; Boeynaems et al., 2017). The arginine DPRs interact with the LCD of proteins involved in the formation of membraneless organelles leading to a disruption in their dynamics; although the precise mechanism by which they are toxic to neurons is not yet fully resolved (Moens et al., 2017). At present it has not been demonstrated whether poly-GR or poly-PR is sufficient or necessary to cause the pathology associated with C9ALS/FTD; i.e. TDP-43 proteinopathy. The development of murine

models overexpressing the arginine DPRs will be important to this. Human post-mortem data suggests the DPRs precede TDP-43 pathology; however, DPR inclusions do not correlate with regional neurodegeneration, whereas TDP-43 burden does (Mackenzie et al., 2014). The arginine DPRs are however less aggregation prone due to their hydrophilic properties, hence sensitive assays to measure soluble forms of the arginine DPRs will be important in helping clear up the current mismatch between model systems and post-mortem analysis.

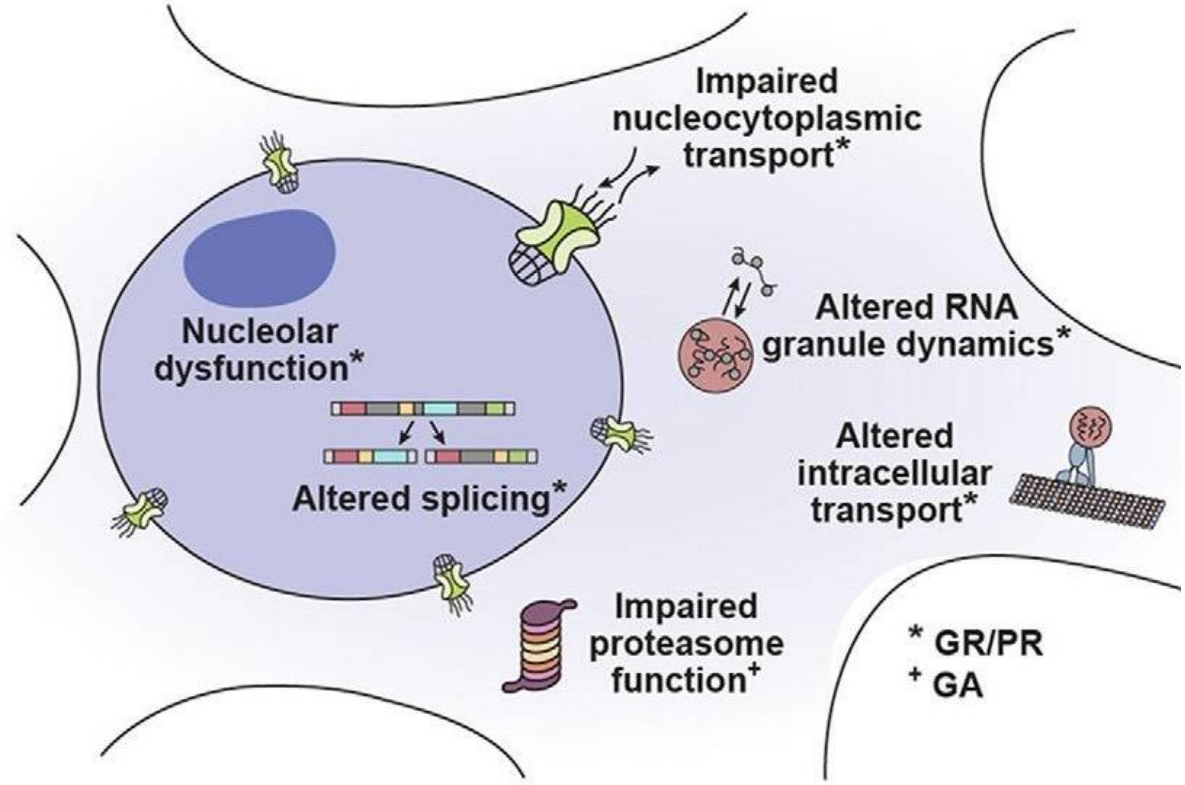

Figure 1.21 | Cellular impairments associated with DPR toxicity. Both poly-GR/PR and poly-GA have shown to be toxic in different model systems. Cellular functions associated with poly-GR/PR toxicity include nucleolar dysfunction, altered splicing, impaired nucleocytoplasmic transport, perturbed RNA granule dynamics and altered intracellular transport. Poly-GA has been shown to impair the function of the proteasome. Adapted from Friebaum & Taylor, 2017.

1.11 Loss vs gain of function

Cellular and animal models of the G_4C_2 repeat expansion have provided major insights into the C9ALS/FTD disease mechanisms. Knockout of the *C. elegans* (Therrien et al., 2013) and Zebrafish orthologus of *C9orf72* (Ciura et al., 2013) has led to motor phenotypes and neurodegeneration. However, a common finding from all murine knockout models is the lack of any apparent neurodegeneration or TDP-43 pathology (O'Rourke et al., 2016; Atanasio et al., 2016; Burberry et al., 2016; Sudria-Lopezet al., 2016; Jiang et al., 2016; Sullivan et al., 2016). This discepenacy between models suggests the function of the protein is redundant to some extent and its loss can be compensated in mice and potentially humans too. These findings clearly indicate a loss of protein function is not sufficient to cause the disease. However, the function of C9ORF72 in autophagy, immune dysregulation, axon growth and stress granule dynamics propose a modulatory role for a loss of protein function in disease pathogensis. Indeed, it has recently been demonstrated by Shi et al., (2018) that repeat expansion associated reductions in C9ORF72 levels can cause neurodegeneration via an accumulation of glutamate receptors and defective clearance of the DPRs. This is not entirely surprising considering the function of C9ORF72 in autophagy and that the DPRs are mainily processed through autophagy (Cristofani et al., 2018), hence a loss of protein function in disease may make neurons hypersensitive to DPR toxicity. One potential explanation as to why C9ORF72 knockout mice show no overt neurodegenerative phenotype may be lack due to the gain of function pathology in these models - if C9ORF72 modulates neuronal vulnerability to toxic stimuli rather than its loss causeing toxicity directly. For example, DPRs cause a mis-splicing of the *EAAT2* transporter in astrocytes (Kwon et al., 2014) which may potentially cause excitotoxicty in neuronal populations with an accumulation of glutamate receptors due to reduced levels of C9ORF72 (Shi et al., 2018). Such findings suggest increasing/restoring C9ORF72 protein amounts may help modulate gain of function related toxicity.

G_4C_2 repeat expansion models provide clear evidence that the toxic gain of function mechanisms are both necessary and sufficient to cause neurodegeneration (Chew et al., 2015), although they are limited by their inability to distinguish between RNA and DPR toxicity. To circumvent this several studies have developed strategies to study

the contribution of repeat RNA and DPRs separately. The consensus from such work is that the DPRs, particularly the arginine DPRs, are highly toxic whereas the toxicity of the G_4C_2 RNA in insolation is milder – with its toxicity depending on the subcellular localisation and model system in which the repeat RNA is expressed. *C9orf72* BAC transgenic mice present with RNA foci but no neurodegeneration (O'Rourke et al., 2015; Peters et al., 2015). In *Drosophila* Moens et al., (2018) observed no neurodegenerative phenotypes in flies expressing either sense of anti-sense repeat RNA despite the presence of both cytoplasmic and nuclear foci in neurons. In *Zebrafish* however both sense and anti-sense repeat RNAs induce toxicity (Swinnen et al., 2018). The discrepancies between model systems using similar constructs is puzzling. One potential explanation for this is that RNA binding proteins essential for repeat RNA toxicity are not conserved in *Drosophila*. However, Moens et al., (2018) show *Glorund*, the *Drosophila* homologue of hnRNP H is sequestered into RNA foci in flies and when knocked down via RNAi causes a degenerative phenotype in flies, suggesting the sequestration of RNA binding proteins into RNA foci is not sufficient to cause neurodegeneration. Indeed Swinnen et al., (2018) found no direct correlation between RNA foci and toxicity in their *Zebrafish* model, however the authors note the RNA was mainly localised to the cytoplasm, both diffisuley as well as in foci – suggesting RNA toxicity may result from cytoplasmic, rather than nuclear repeat RNA. ISwinnen et al., (2018) report that Pur-alpha overexpression reduces RNA foci numbers without being present in foci, suggesting it interacts with soluble repeat RNA, which may drive repeat RNA toxicity. Indeed, a possible toxic effect of cytoplasmic RNA not incorporated into foci is supported by neuronal patient iPSCs (Burguete et al., 2015). Therefore, one potential explanation for lack of any phenotype observed following RNA expression in flies could be the marked RNA foci in this model when compared to the more diffuse cytoplasmic RNA seen in the *Zebrafish* model.

Protein only constructs designed to study the individual DPRs in isolation of the repeat RNA have given a broad consensus that the arginine DPRs (poly-GR and poly-PR) are highly toxic (Moens et al., 2017), indeed it appears they are the 'primary' gain of function mechanism. The disruption of membraneless organelles that result from abnormal phase transitions induced by these DPRs can account for many of the widespread cellular abnormalities observed in C9ALS/FTD models. Of

the other DPRs only poly-GA has shown mild toxicity depending on the model system used, with *Drosophila* being resisent to poly-GA toxicity (Mizielinska et al., 2014) whereas neurodegeneration, cognitive dysfunction and motor impairment is seen following poly-GA expression in mice (Zhang et al., 2016). The development of rodent models to study arginine DPR toxicity will be vital to understanding the contribution of these DPRs to neurodegeneration and ALS/FTD releated phenotypes. An important issue with DPR toxicity is the lack of spatial correlation between DPRs and regions that undergo severe cell loss in FTLD and ALS in addition to their high abundance in areas of the brain not typically associated with disease such as the cerebellum. Although, a more recent thorough and standardised analysis of DPRs in clinically-related and clinically-unrelated regions of the brain found poly-GR was more abundant in clinically related than clinically unrelated areas (Saberi et al., 2018). The lack of association in previous studies may owe to arginine methylation, which would lead to several complex post-translational modifications of poly-GR. It possible different poly-GR antibodies used in different papers detect different species. Decoupling RNA and DPR toxicity remains however a challenge without creating artificial constructs or relying on overexpression systems. It it is clear both can be toxic (particularly for the arginine DPRs) in a dose-dependent manner, however whether they are toxic at physiological levels is still unknown. A better understanding of RAN translation and the mechanisms underpinning this non-conventional form of translation will enable the development of tools which will allow DPR expression, translated at physiological levels, to be turned off and on in endogenous systems such as iPSCs without disrupting the structure of the repeat RNA (figure 1.22). Such a tool would also help establish a pathogenic chain of events in C9ALS/FTD which at present is lacking. Another strategy would be to develop repeat RNA constructs in which a specific DPR (such as poly-GR) is produced in frame with a degradable reporter gene which would allow the expression of that DPR to be modulated without disrupting translation of the other DPRs. Current DPR only mouse models (Zhang et al., 2016; Zhang et al., 2018) show neurodegeneration and behavioural phenotypes but no TDP-43 pathology – where as BAC transgenic mice with a combination of both DPR and RNA pathology (Chew et al., 2015; Lui et al., 2016) show TDP-43 inclusions and neurodegeneration. Such data argues for a combination of both the gain of function mechanisms to be fully required to recapitulate disease pathogenesis.

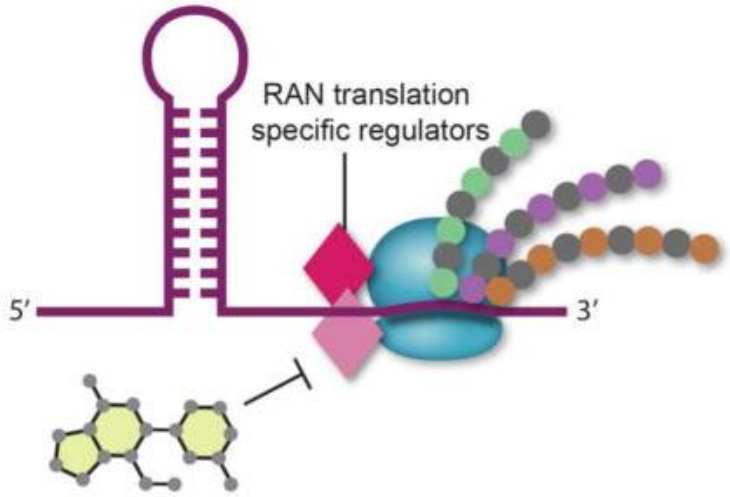

Figure 1.22 | Further work needed to study DPR toxicity at physiological levels. To separate RNA and DPR toxicity without using overexpression models and to study DPR toxicity at physiological levels the identification of compounds that can specifically block RAN translation will be needed. The development of such compounds requires a greater understanding of RAN translation mechanisms and proteins that specifically regulate RAN translation. Adapted from Gitler and Tsuiji, 2016.

1.12 Summary of disease mechanisms

In summary major progress has been made since the 2011 discovery of the repeat expansion in *C9ORF72* in elucidating the underlying pathogenic mechanisms. There is evidence to support both loss and gain of function mechanisms; however current research demonstrates a loss of C9ORF72 protein function is not sufficient to cause disease. Both the G_4C_2 RNA and DPR expression (in particular, poly-GR and poly-PR) have been shown to be neurotoxic and disrupt a number of cellular processes. Establishing which of these correlate with neuronal loss will be an important future area of work. It is likely both the gain and loss of function mechanisms act synergistically to induce disease, indeed both the repeat RNA and arginine DPRs have been shown to disrupt nucleocytoplasmic transport and cause nucleolar dysfunction. A loss of protein function may also play a modulatory role via its role in autophagy, nucleocytoplasmic transport and stress granule formation.

1.13 TDP-43 dysfunction in C9ALS/FTD

C9ALS/FTD, like 97% of all ALS cases and 45% of FTD cases is a TDP-43 proteinopathy (Ling et al., 2013). Unlike DPR and RNA foci pathology, TDP-43 pathology is highly correlated with both brain areas showing the most neurodegeneration and clinical symptoms (Mackenzie et al., 2014; DeJesus-Hernandez et al., 2017). Based on this it has been proposed TDP-43 mislocalisation and aggregation is the most likely effector of C9 repeat associated toxicity; however no casual relation and mechanistic link has been fully established (Edbauer and Haass, 2016). Edbauer and Haass, (2016) note pathogenic TDP-43 mutations cause mainly regional TDP-43 pathology highly similar to the TDP-43 pathology seen in sporadic ALS/FTD cases; indicating the frontotemporal cortex and motor neurons are prone to cytoplasmic TDP-43 accumulation.

A multiple hit hypothesis has been proposed by Dormann and Haass, (2011) in which multiple stressors are required for the accumulation of TDP-43 into insoluble inclusions. The first hit is argued to be defective nuclear import which causes a cytoplasmic accumulation of TDP-43 (Dormann and Haass, 2011); a second hit is required for aggregate formation which is suggested to be a cellular stress that causes stress granule formation (Dormann & Haass, 2011). Such a multiple-hit model has been demonstrated for FUS; FUS mutations that disrupt the NLS lead to a diffuse cytoplasmic distribution of FUS without the formation of FUS aggregates. However, when cells are stressed cytoplasmic FUS begins to accumulate in SGs (Bosco et al., 2010; Dormann et al., 2010). Edbauer and Haass, (2016) propose that the DPR proteins and/or the repeat RNA are the additional stressors that cause TDP-43 pathology in C9/ALS FTD. Edbauer and Haass, (2016) have proposed an amyloid cascade model for *C9ORF72* pathogenesis in which the DPRs and/or repeat RNA trigger a toxic cascade prior to the onset of clinical symptoms. The authors note this is akin to what occurs in AD where accumulation of the Aβ is the most probable initial trigger at the start of the disease cascade, followed by subtle synaptic and early cognitive dysfunction, the deposition of tau into tangles leading to the impairment of synaptic functional and neuronal loss. Importantly Edbauer and Haass, (2016) note in this cascade the Aβ pathology occurs decades before dementia onset in the neocortex and then spreads centripetally toward the brain

stem; Edbauer and Haass, (2016) argue by the time AD is diagnosed the Aβ pathology has plateaued and no longer correlates well with the progression of AD as is the case with the DPRs in C9ALS/FTD. Using AD as an example to model C9ALS/FTD pathogenesis Edbauer and Haass (2016) propose neurodegeneration is caused by the cytoplasmic accumulation of abnormally phosphorylated TDP-43 similar to abnormally phosphorylated tau in AD. Hence, they argue the lack of any association between the DPRs and neurodegeneration is potentially explained by the triggering of a slow cascade that occurs before clinical symptoms which later sustains itself independently of the trigger like what has been proposed for Aβ plaques in AD.

In support of this post-mortem analyses in patients with the *C9ORF72* expansion who have died early of the disease show abundant DPR pathology only; further indicating DPR pathology precedes that of TDP-43 (Proudfoot et al 2014; Baborie et al 2015; Davidson et al 2016). Vatsavayai et al. (2016) could show TDP-43 pathology occurs only at symptomatic and not pre-symptomatic stages of disease; DPRs were however found at both stages. Cellular studies have also linked a loss of C9ORF72 protein function to TDP-43 accumulation. For example, Sellier et al. (2016) found reduced expression of C9ORF72 via shRNA mediated knockdown in mouse embryonic cortical neurons caused a ~30kDA fragment of mouse TDP-43 (Tdp-43) to accumulate. Phosphorylated cytoplasmic aggregates of Tdp-43 were also observed in these neurons. Furthermore, overexpression of C9ORF72 reduced the aggregation propensity of the D196G mutant of TDP-43. Other cellular models have not replicated this; Maharjan et al. (2016) saw TDP-43 levels and localisation were unchanged by either C9-CRISPR knockdown or overexpression of the long isoform of C9ORF72. Xiao et al. (2015a) did observe a depletion of C9-S from the nuclear membrane in patient spinal cord motor neurons, further this loss from the nuclear membrane correlated with a mislocalisation of KPNB1, RAN and TDP-43. However, no animal model of C9ORF72 loss of function has shown any sort of TDP-43 pathology (Ciura et al., 2013; Koppers et al., 2015; O'Rourke et al., 2016; Atanasio et al., 2016; Burberry et al., 2016; Sudria-Lopezet al., 2016; Jiang et al., 2016; Sullivan et al., 2016). This strongly argues against a loss of function of the protein being sufficient to cause TDP-43 accumulation.

If the DPRs and/or RNA foci act as such stressors, causing TDP-43 dysfunction leading to subsequent neurodegeneration as is proposed Edbauer and Haass., (2016), the question is which of these gain of function pathogenic species generated from the G_4C_2 is the main stressor in this cascade. For example, are both the DPRs and pathogenic RNA necessary for TDP-43 dysfunction, or is just one of these sufficient, and if so, in the case of the DPRs, is one or all of the DPRs required for TDP-43 cytosolic accumulation and aggregation. Murine C9 bacterial artificial chromosome (BAC) models have been generated that produce RNA foci and RAN dipeptides early during the life of these mice, but show no TDP-43 pathology and no neurodegeneration or motor phenotypes (O'Rourke et al 2015; Peters et al 2015). Neurodegeneration is seen in murine models that present TDP-43 pathology (Chew et al 2015; Liu et al 2016a); interestingly in the Liu et al. (2016a) C9 BAC model TDP-43 aggregation occurred at end-stage, after the formation of both DPRs and RNA foci. In the Chew et al. (2015) viral vector model, of the 250 cortical cells with TDP-43 inclusions all contained at least one nuclear sense RNA foci; whilst ~75% of cells also contained a poly-GA aggregate. Chew et al. (2015) suggest as all cells investigated harboured RNA foci, that the G_4C_2 RNA and/or the RNA foci may be the instigator of TDP-43 pathology. Cooper-Knock et al. (2015) also suggest the anti-sense C_4G_2 RNA may drive TDP-43 pathology; finding that anti-sense RNA foci, but not sense RNA foci, correlate with TDP-43 mislocalisation in patient motor neurons. In a reported RNA-only pure repeat *Drosophila* generated by Xu et al. (2013) increased cytoplasmic *Drosophila* TDP-43 is observed (Zhang et al., 2015); although undetectable levels of DPRs cannot be ruled out in this model; and as noted previously are detected under certain conditions. Jiang et al. (2016) in their BAC model observed increased phosphorylated TDP-43 at end stage but saw no mislocalisation or aggregation; both sense and anti-sense foci were produced in these mice in addition to all three sense DPRs. The anti-sense DPRs were not detectable in this model and were also not probed for by Liu et al. (2016a).

A number of studies suggest it is the DPRs that cause TDP-43 pathology. Liu et al. (2016a) observed TDP-43 inclusions and behavioural phenotypes only in BAC mice that produced DPRs. Khosravi et al., (2017) suggest poly-GA may mediate TDP-43 mislocalisation. Khosravi et al., (2017) generated a TDP-43 NLS reporter which consisted of RFP fused to the NLS of TDP-43. In control conditions the TDP-43 NLS

reporter was completely nuclear. However, when co-expressed with poly-GA significant levels of the reporter were found in the cytoplasm indicating nuclear import mediated by the TDP-43 NLS is perturbed. Manual counting of cells revealed ~40% of cells containing a poly-GA inclusion showed mislocalisation of the reporter. Manual counting also revealed ~20% of poly-GR and poly-PR inclusion bearing cells also had mislocalisation of the TDP-43 NLS reporter. Manual analysis revealed poly-GR led to a significant mislocalisation of the reporter although this significance was lost when an automated analysis was performed. It must be noted the authors did not comment on the reported frequency of poly-GR inclusions which could bias the results; poly-GR inclusions are rarer than poly-GA inclusions (Schludi et al., 2015) potentially due to the high toxicity of the peptide, especially in model systems (Gitler and Tsuiji, 2016; Moens et al., 2017) and its flexible-coil secondary structure (Freibaum and Taylor, 2017). Additionally, the authors did note that expression levels of poly-PR were significantly lower than poly-GA.

Khosravi et al. (2017) also found importin α/β-mediated nuclear import, the import pathway involved in TDP-43 shuttling, is perturbed by poly-GA expression in addition to also enhancing the localisation of TDP-43 to cytoplasmic granules in transfected hippocampal neurons. No change in the amount of TDP-43 in cytoplasmic granules was seen for poly-GR; a significant increase was seen for poly-PR from a manual analysis but after automated analysis the significance was lost. Interestingly forcing poly-GA into the nucleus prevented cytoplasmic TDP-43 mislocalisation. Finally, Khosravi et al. (2017) observed overexpression of proteins involved in importin α/β-mediated shuttling of TDP-43 such as KPNA3, KPNA4 and Nup62, but not Ran or RanGAP1, restored nuclear localisation of the TDP-43-NLS reporter following poly-GA expression. It is important to point out that the authors saw no changes in the localisation of endogenous TDP-43 following DPR expression. Lee et al. (2017) observed significant TDP-43 cleavage in poly-GA expressing cells in a dose responsive manner which increased in parallel to the appearance of higher molecular weight poly-GA on western blot. However, Lee at al. (2017) did not report any TDP-43 aggregates or mislocalisation in poly-GA expressing cells. However, in poly-GA only mouse models developed by Zhang et al. (2016) only very rare phosphorylated TDP-43 aggregates were seen despite neurodegeneration, behavioural abnormalities and evidence for impaired nuclear import. Schludi et al.

(2017) in their poly-GA mouse model did observe phosphorylated TDP-43 but again no TDP-43 aggregates or even cytoplasmic accumulation of TDP-43 at 12 months of age.

Other studies suggest the highly toxic arginine DPRs cause TDP-43 mislocalisation. A cell culture model expressing 80 G_4C_2 repeats developed by Mori et al. (2016) demonstrated cytoplasmic TDP-43 re-distribution when all three sense DPRs, poly-GP, poly-GA and poly-GR were present in cells, but not with either just poly-GP or poly-GA or the combination of the two (Mori et al 2016). The authors note in the replies to one of the reviewers of their paper that they do indeed speculate poly-GR accumulation contributes to TDP-43 mislocalisation; however, they were unable to quantify this as cells expressing poly-GR only were too rare to find. Furthermore, poly-GA and poly-GR double-positive cells, but not cells with only poly-GA, frequently had cytosolic TDP-43 accumulation. Indeed, the lack of poly-GR only cells may be due to the high toxicity of the peptide; but when present with poly-GA its detrimental effects may potentially be modulated, but they are still sufficient to disrupt TDP-43 dysfunction. Such events are consistent with the finding by Yang et al. (2015) that poly-GA sequesters poly-GR and modulates its toxicity in *Drosophila* and cells. Indeed Yang et al. (2015) hypothesise poly-GA may be neuroprotective during the early stages of disease by trapping the highly toxic poly-GR in inclusions allowing it to be efficiently degraded; however, over time elevated levels of non-aggregated poly-GR will increase to the point where it starts to disrupt cellular pathways. A recent analysis of DPRs in clinically-related and clinically-unrelated regions of the CNS found poly-GR was more abundant in clinically related compared clinically unrelated areas and co-localised with TDP-43 aggregates found in dendrites, such findings were not observed for any of the other DPRs (Saberi et al., 2018).

Based on their mechanisms of toxicity the arginine DPRs poly-PR and poly-GR seem likely candidates to cause TDP-43 dysfunction. As discussed, the arginine DPRs disrupt nucleocytoplasmic transport; including proteins involved in the importin α/β-mediated pathway involved in the shuttling of TDP-43 between the cytoplasm and nucleus (Jovičić et al., 2015; Boeynaems et al., 2016a; Lee et al., 2016; Shi et al., 2017). Boeynaems et al. (2016b) argue based on a review of the literature and bioinformatics analysis that deficits in nucleocytoplasmic transport play a key initial

role in the initiation of disease. Dormann and Haass, (2011) propose in their multiple hit hypotheses for TDP-43 aggregation that the first hit is defective nuclear import, which leads to cytoplasmic accumulation of TDP-43. The arginine DPRs provide this first hit; however, Dormann and Haass, (2011) argue without a second hit the protein remains soluble and no aggregates will be observed. Dormann and Haass, (2011) propose impaired SG dynamics may be the key second hit. The arginine-rich DPRs however provide this important second hit through their disruption of the formation of membraneless organelles (Lee et al., 2016; Lin et al., 2016; Boeynaems et al., 2017). SGs dynamics are particularly interesting in the context of TDP-43 cytoplasmic accumulation and aggregation. SGs are thought to be an important step in the aggregation of RNA binding proteins in ALS and FTD (Dormann and Haass., 2011). SGs are cytoplasmic membrane less organelles that form in response to cellular stress (Aulas and Vande Velde., 2015).

The transformation of liquid droplets like SGs into pathological aggregates is an attractive hypothesis via which arginine-DPR mediated TDP-43 aggregation may occur in C9ALS/FTD. Indeed, several groups have proposed that the accumulation of TDP-43 into cytoplasmic SGs leads to its aggregation (Ramaswami et al, 2013; Li et al., 2013; Aulas and Vande Velde, 2015 Becker et al., 2017). TDP-43 is a known component of stress granules; previous work has shown it forms part of stress granules by interacting with stress granule proteins including TIA-1 (Liu-Yesucevitz et al., 2010; Bentmann et al., 2012). Indeed, there is good evidence to show pathological TDP-43 inclusions may originate from SGs. Becker et al. (2017) investigated the role of the RNA binding protein ataxin-2 in TDP-43 aggregation. Ataxin-2 is suppressor of TDP-43 toxicity in flies and yeast and intermediate-length polyglutamine expansions in this gene are a risk factor ALS (Elden et al., 2010). Furthermore ataxin-2 regulates the assembly of SGs (Kaehler et al., 2012). Hence Becker et al. (2017) found ataxin-2 knockdown significantly decreased the number of SGs with endogenous TDP-43. *TDP-43*$^{Tg/Tg}$*Atxn2*$^{-/-}$ transgenic mice had markedly decreased TDP-43 aggregation supporting the hypothesis that decreasing ataxin-2 reduces the association of TDP-43 with SGs and its propensity to aggregate.

Recently Mackenzie et al. (2017) discovered mutations in the PrLD of the RNA binding SG protein TIA1 cause ALS/FTD. Importantly TIA1 ALS/FTD was a TDP-43 proteinopathy. TIA1 mutations lowered the protein concentration required for the

protein to undergo LLPS and form liquid droplets at higher protein concentrations; hence TIA1 mutant phase separation is enhanced by stronger intermolecular protein-protein interactions – similar to what is proposed for the increased LLPS caused by poly-GR and poly-PR (Lee et al., 2016). TIA mutations led to SGs which had impaired disassembly; importantly TDP-43 containing SGs accumulated because of TIA mutations. The recruitment of TDP-43 into these SGs led to TDP-43 becoming completely immobile and insoluble. Normally the solubility change in TDP-43 was reversible 3 hours after 30 minutes of stress; however, when stress persisted for 1 hour TDP-43 became permanently insoluble. This supports the mounting evidence that perturbed SG dynamics are involved in pathological TDP-43 aggregation in ALS/FTD.

The finding that poly-GR and poly-GR perturbs phase transitions disrupting both nucleocytoplasmic transport and SG formation provides the molecular basis for TDP-43 aggregation in C9ALS/FTD. An important step in investigating the development of TDP-43 pathology in C9ALS/FTD will be the study of TDP-43 in poly-GR and poly-PR animal models which is yet to be done. The investigation of TDP-43 localisation in animal models expressing these DPRs from alternative codons will be important step in understanding if these DPRs are indeed sufficient to cause TDP-43 proteinopathy.

1.14 Aims

A non-coding G_4C_2 repeat expansion in C9ORF72 has been identified as major genetic cause of ALS and FTD and the combined ALS/FTD. An amyloid cascade hypothesis for C9 mediated neurodegeneration has been proposed in which stressors causes TDP-43 aggregation and cytoplasmic mislocalisation (Edbauer and Haass, 2016). It has been proposed that the DPRs and/or RNA foci act as such stressors, causing TDP-43 dysfunction leading to subsequent neurodegeneration (Edbauer and Haass, 2016). Hence this project aims to establish whether there is a link between the DPRs and TDP-43 pathology. The three major aims are as follows:

1. Characterise RAN translation and the DPRs produced in *Drosophila* harbouring uninterrupted G_4C_2 repeats of different lengths.

2. Generate and validate *Drosophila* expressing the DPRs from non G_4C_2 codons.

3. Investigate the potential pathogenic link between the DPRs and TDP-43 mislocalisation in these alternative codon flies.

Chapter 2: Materials and Methods

2.1 Drosophila husbandry

2.1.1 Standard Fly Storage and food

Fly stocks were maintained at 25ºC in humid incubator (LMS) set to a 12 hour light/dark cycle. Flies were maintained on standard cornmeal food.

Recipe for standard cornmeal food (per 1L):

Distilled H_2O – 1L

Dried active yeast (Allinson) – 17.5g

Soya flour (Neal's Yard) – 10g

Sugar (Tate and Lyle) – 32g

Fine cornmeal (Dunn's River) – 73.1g

Malt extract (Oxoid) – 46.2g

Agar (Sigma) - 6g

Nipagin (Methyl 4-hydroxybenzoate) (Sigma) diluted in 96% v/v ethanol (Fisher) – 2.5g in 25ml ethanol

Propionic acid (Aldrich) – 5ml

2.1.2 Fly Storage for ageing experiments and ageing food

For aging experiments day 1 male and female flies were collected and allowed to mate for ~48 hours on standard cornmeal food. After ~48 hours the males were removed, and the females were transferred to and maintained on a 15% sugar/yeast medium at 25ºC in humid incubator (LMS) set to a 12 hour light/dark cycle. Ageing flies were flipped onto new food every 3-4 days.

Recipe for 15% sugar/yeast medium (per 1L):

Distilled H_2O – 1L

Dried active yeast (Allinson) – 150g

Sugar (Tate and Lyle) – 150g

Agar (Sigma) - 15g

Nipagin (Methyl 4-hydroxybenzoate) (Sigma) diluted in 96% v/v ethanol (Fisher) – 3g in 30ml ethanol

Propionic acid (Aldrich) – 3ml

2.2 GAL4/UAS

All fly lines were obtained from the Bloomington Stock Centre unless stated otherwise

The GAL4/UAS system allows both selective expression and knockdown of target genes in a cell and tissue specific manner (figure 2.1). The system consists of two parts; the yeast transcription factor GAL4 and the upstream activating sequence (UAS). The GAL4 binds to UAS to initiate transcription of a gene of interest downstream of the UAS. GAL4 expression is directed by a tissue specific native promoter; hence the gene of interest is only transcribed where the GAL4 is present - in a defined population of cells. The key feature of this tool is that the GAL4 gene and the UAS-target are separated into two different transgenic lines, which when crossed brings the GAL4 and UAS together allowing selective overexpression of a gene of interest in the tissue/cell type of choice (St Johnston, 2002).

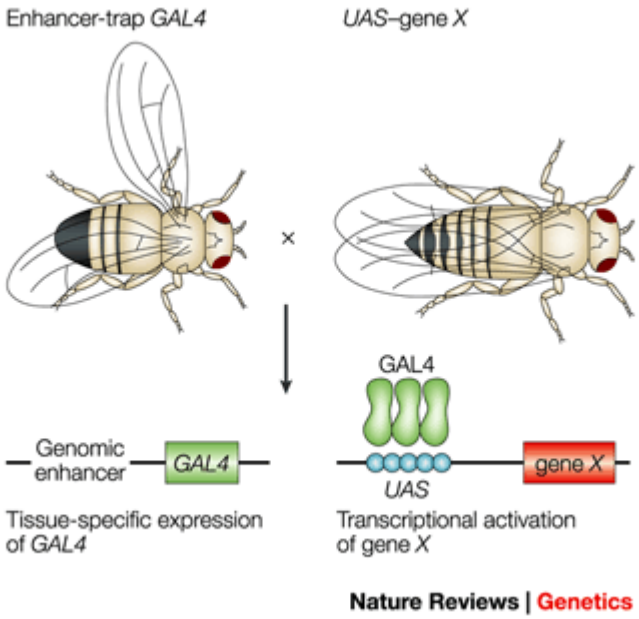

Figure 2.1 | The GAL4/UAS system. The GAL4/UAS system allows tissue or cell specific expression of a gene of interest. The yeast transcription factor GAL4 recognises the upstream activating sequences (UAS). The gene of interest is cloned downstream of the UAS promoter. GAL4 binds to UAS enabling transcription of the gene of interest. GAL4 expression is under the control of a tissue specific enhancer, hence the gene of interest is only expressed where GAL4 is selectively expressed. The GAL4 and UAS-target gene inserts are kept separate in two different transgenic lines. Only by crossing these two fly lines will the GAL4/UAS system be activated. From St Johnstone, (2002).

2.2.1 GAL4 Lines

For the present study the *Elav^{c155}-GAL4* driver (BL#458) was chosen for selective overexpression of a gene of interest in all differentiating neurons. Expression of transgenes in ellipsoid body neurons was achieved using *UAS-mCD8-GFP;EB1-GAL4* (generated in Hirth lab from BL#5137 and BL#432). Overexpression in photoreceptors was achieved using *GMR-GAL4* (BL#1104) and *Rhodopsin 1-GAL4* (*RH1-GAL4*) (gift from Manolis Fanto - King's College London). Salivary gland expression studies were carried out with *FKH-GAL4* (gift from Eric Baehrecke - University of Massachusetts).

2.2.2 UAS Lines

The G_4C_2 fly lines used in this study were generated by Alan Stepto and Yoshitsugu Adachi. Briefly, UAS constructs were designed in which different lengths of G_4C_2

repeats were cloned into the 3'UTR of disease unrelated gene DsRed2. All G_4C_2 constructs were inserted into the genome at a common insertion site via φC31-mediated site-specific integration into flies harbouring an attP insertion at site ZH-86FB on the third chromosome (BL#23647). The following lines were generated:

w^{1118}; +; UAS-DsRed2

w^{1118}; +; UAS-DsRed2-(G4C2)8A

w^{1118}; +; UAS-DsRed2-(G4C2)32

w^{1118}; +; UAS-DsRed2-(G4C2)56

w^{1118}; +; UAS-DsRed2-(G4C2)64

w^{1118}; +; UAS-DsRed2-(G4C2)128A

w^{1118}; +; UAS-DsRed2-(G4C2)128B

w^{1118}; +; UAS-DsRed2-(G4C2)8B

w^{1118}; +; UAS-DsRed2-(G4C2)38

Generation of transgenic flies was performed by the Cambridge transgenic fly facility (Cambridge University, UK) for the 8B and 38 G_4C_2 constructs. Due to facility closing after the generation of 8B and 38 transgenic lines, the remaining UAS-G_4C_2 constructs were sent to Bestgene, USA for genomic insertion.

Gene fragments coding for individual DPR sequences were generated by GeneArt (ThermoFisher Scientific). Fragments were ligated into the pUAST*attB (*Bischof et al., 2007) also by GeneArt (ThermoFisher Scientific). UAS-DPR Constructs were then sent to Bestgene, USA for transgenesis. All UAS-DPR constructs were inserted into the genome at a common insertion site via φC31-mediated site-specific integration into flies harbouring an attP insertion at site attp40 on the second chromosome (y1 w67c23; P{CaryP}attP40). Genomic PCR confirmation of insertion onto the attp40 landing site was confirmed by BestGene, USA. The following lines were generated:

w^{1118}; UAS-GA8; +

w^{1118}; UAS-GA64; +

w^{1118}; UAS-GR8; +

w^{1118}; UAS-GR64; +

w^{1118}; UAS-PR8; +

w^{1118}; UAS-PR64; +

w^{1118}; UAS-PA8; +

w^{1118}; UAS-PA64; +

The *UAS-288 RNA-Only* line was kindly provided by Adrian Isaacs and Linda Partridge (University College London). Construction of genomic TBPH is described in a previous publication (Diaper et al., 2013a). Nuclear localisation signal (NLS) mutant (gΔNLS) was generated via mutation containing PCR by Yoshitsugu Adachi (Hirth lab, KCL).

2.3 Generation of polyclonal and monoclonal poly-GP antibodies

Two anti-GP antibodies were generated during this project.

An anti-GP rabbit polyclonal antibody was generated by immunisation of rabbits with the following peptide sequence – GPGPGPGPGPGPGPGP (GPx8). The resulting antiserum was purified against the immunisation peptide. This work was conducted by Invitrogen.

Anti-GP mouse monoclonal was generated by immunisation of mice with the following peptide sequence - GPGPGPGPGPGPGPGPGPGP (GPx10). The resulting antiserum was purified against the immunisation peptide. This work was conducted by Abmart, China.

2.4 Western blotting

2.4.1 Solutions for western blotting for Drosophila tissue

Radioimmunoprecipitation assay (RIPA) buffer and RIPA+ buffer for Drosophila tissue

50 mM Tris (Roche) with HCL added (Sigma) to achieve a pH of 7.5
150 mM NaCl (Fisher)
1% (v/v) NP-40 (Calibochem)
5 mM EDTA (Sigma)
0.5% (w/v) Sodium deoxycholate (Sigma)
0.1% (w/v) SDS (Sigma-Aldirch)
Made up to 100ml with ddH_2O
To make RIPA+ one complete mini protease inhibitor tablet (Roche) and one phosphatase inhibitor tablet (Roche) was dissolved in 7ml of RIPA buffer.

Laemmli sample buffer (5x)

312.5 mM Tris (Roche) with HCL (Sigma) added to achieve a pH of 6.8
10% (w/v) SDS (Sigma-Aldirch)
250 mM DTT (Life Technologies)
50% (v/v) Glycerol (Santa Cruz)

Resolving gel

Resolving gel buffer:
1.5M Tris (Roche) with HCL (Sigma) added to achieve a pH of 6.8
Made up to desired volume with ddH_2O
Resolving gel:
30% degassed acrylamide (% v/v varied depending on gel %)
ddH_2O (% v/v varied depending on gel %)
25% (v/v) resolving gel buffer
1% (v/v) 10% (w/v) SDS

Stacking gel

Resolving gel buffer:

0.5M Tris (Roche) with HCL (Sigma) added to achieve a pH of 6.8

Made up to desired volume with ddH_2O

Stacking gel:

30% degassed acrylamide (% v/v varied depending on gel %)

ddH_2O (% v/v varied depending on gel %)

25% (v/v) stacking gel buffer

1% (v/v) 10% (w/v) SDS

Running buffer (1x)

25 mM Tris (Roche)

192 mM Glycine (Sigma)

0.1% (w/v) SDS (Sigma-Aldrich)

Made up to 1L with ddH_2O

Transfer buffer (1x)

25 mM Tris (Roche)

192 mM Glycine (Sigma)

20% (v/v) Methanol (Fisher)

Made up to 1L with ddH_2O

Phosphate buffered saline (PBS)

5X PBS tablets (Sigma) dissolved in 1L ddH_2O

Phosphate buffered saline-Tween (PBST)

5X PBS with 0.1% (v/v) Tween-20 (Sigma)

2.4.2 Solutions for western blotting from human post mortem brain tissue

Radioimmunoprecipitation assay (RIPA) buffer and RIPA+ buffer for human post mortem tissue

50 mM Tris (Roche) with HCL added (Sigma) to achieve a pH of 7.5

150 mM NaCl (Fisher)

1% (v/v) NP-40 (Calibochem)

5 mM EDTA (Sigma)

0.5% (w/v) Sodium deoxycholate (Sigma)

0.1% (w/v) SDS (Sigma-Aldirch)

Made up to 100ml with ddH_2O

To make RIPA+ one complete mini protease inhibitor tablet (Roche) was dissolved in 7ml of RIPA buffer.

UREA lysis buffer

7M urea (Sigma)

2M thiourea (Sigma)

4% CHAPS (CALBIOCHEM)

30 mM Tris (Sigma)

Made up to desired volume with ddH_2O

Laemmli sample buffer (5x), resolving gel, stacking gel, running buffer (1x), transfer buffer (1x), PBS and PBST used in the western blotting lysate from brain is identical to that used for Drosophila *tissue*.

2.4.3 Protein extraction for Drosophila samples

For *Drosophila* samples, mated females were snap frozen in liquid nitrogen and then vortexed to remove the heads. Heads were homogenised using a sterile pestle in RIPA+ buffer at a volume 40ul per 6 fly heads. Following homogenisation samples were sonicated and then left on ice for 40 minutes. The homogenate was then spun at 14,000rpm at 4ºC and the supernatant removed.

2.4.4 Protein extraction from human brain

2.4.4.1 Extraction of soluble protein from human brain tissue

100mg of frontal cortex tissue was provided by the MRC London Neurodegenerative Diseases Brain Bank, case details can be found in appendix 9; table A9.1. Frontal cortex tissue was homogenized using 15ml falcon tubes (Fisher Scientific) in 1ml of RIPA+ buffer using the Bio-Gen PRO200 Homogenizer (PRO Scientific Inc). The falcon tubes were then spun at 4000g for 10 minutes at 4°C. The pellet consisting of unhomogenised tissue and blood was discarded and the supernatant removed. The supernatant was transferred to 1.5ml RNase free Microfuge tubes (Ambion) and

spun at 14,000rpm for 10 minutes at 4°C. The supernatant was removed, and this was taken as the soluble fraction. The remaining pellets was frozen and stored for later extraction of insoluble protein.

2.4.4.2 Extraction of insoluble protein from human brain tissue

For detection of insoluble KPNA4 in human brain, the pellets described in section 2.4.4.1 were washed four times 10 minutes with RIPA buffer and sonicated. Pellets were then resuspended in 20 µl urea buffer and spun at again at 14,000 rpm for 20 minutes at 4°C. The supernatant was collected and then prepared for western blot.

2.4.5 Determining protein concentration

Protein concentration was determined using the Bicinchoninic kit (BCA) assay kit (Thermo Scientific Pierce). A set of serial dilutions of bovine serum albumin (BSA) (2mg/ml) and RIPA+ buffer was used to produce a standard curve. Lysate was diluted 10x in RIPA+ buffer in duplicate. The assay was performed as per the manufacturer's instructions; the absorption spectra were measured using a spectrophotometer (Perkin Elmer Victor^3). Absorption spectra values were used to determine protein concentration of samples in ug/µl.

2.4.6 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

2.4.6.1 Casting the SDS-PAGE gel

5%, 10% and 15% SDS-PAGE gels were prepared by mixing the following reagents:

Percent Gel	ddH_2O (ml)	30% Degassed Acrylamide/Bis (ml)	Gel Buffer (ml)	10% w/v SDS (ml)
5%	5.7	1.7	2.5	0.1
10%	4.1	3.3	2.5	0.1
15%	2.4	5.0	2.5	0.1

The mixture was degassed for 15 minutes after which 50 µl 10% APS (Sigma) and 5 µl TEMED (Sigma) was added to make the resolving gel. For the stacking gel 50 µl 10% APS and 10 µl TEMED were added.

2.4.6.2 Running the SDS-PAGE gel

Samples were mixed with 5x Laemmli sample buffer to a working concentration of 1x, boiled for 5 minutes at 100 ^{o}C, separated on an SDS-PAGE gel and then transferred to the nitrocellulose membrane (Pall). The membrane was then blocked either in odyssey blocking buffer (OBB) (Li-Cor) (mixed 1:1 in PBS) or 5% dry milk in PBS-T, depending on whether the membrane was being imaged using the Odyssey Fc (Li-Cor) or via chemiluminescent substrate (Amersham ECL Prime Western Blotting Detection Reagent, GE Healthcare). The Primary antibody solution was then left overnight at 4^{o}C. Membranes were then washed 3x 5 minutes in PBS-T and incubated with secondary antibody for 1 hour at room temperature with agitation. Membranes were further washed 4x 5 minutes in PBS-T before imaging. The signal was visualised with either chemiluminescent substrate or by using an Odyssey Fc.

2.4.6.3 Antibodies used for western blotting from Drosophila tissue

The following primary antibodies and dilutions were used;

Antibody	Dilution	Source
Rabbit polyclonal anti-DsRed2	1:1000	Clontech (632496)
Mouse monoclonal anti-GP	1:1000	Abmart (Hirth lab, King's College London)
Mouse monoclonal anti-GA clone 5E9	1:1000	Merck Millipore (MABN889)
Rabbit polyclonal anti-TBPH	1:2000	Hirth lab (King's College London)
Rabbit polyclonal anti-beta actin	1:1000	Abcam (ab8227)
Mouse monoclonal anti-beta tubulin clone E7	1:600	Developmental Studies Hybridoma Bank

The following secondary antibodies and dilutions were used;

Antibody	Dilution	Source
Polyclonal goat anti-rabbit IgG (HandL) conjugated IRDye800	1:10000	Rockland immunochemicals, (611132002)
Polyclonal goat anti-mouse IgG (HandL) conjugated Alexa Fluor 680	1:10000	Thermo Fisher (A21057)

2.4.6.4 Antibodies used for western blotting from human tissue

The following primary antibodies and dilutions were used;

Antibody	Dilution	Source
Goat polyclonal anti-KPNA4	1:750	Abcam (ab6039)
Rabbit polyclonal GAPDH 14C10	1:5000	Cell Signaling Technology (2118S)

The following secondary antibodies and dilutions were used;

Antibody	Dilution	Source
Polyclonal goat anti-rabbit IgG (whole molecule)-Peroxidase	1:10000	Sigma (A0545)
Polyclonal rabbit anti-goat IgG (whole molecule)-Peroxidase	1:10000	Sigma (A8919)

2.5 Dot blotting

2.5.1 Dot blotting protocol for detecting poly-GR

Dot blotting was used for the detection of poly-GR using a protocol previously described by Mizielinska et al. (2014). Fly heads were collected and homogenised in RIPA+ buffer as reported in section 2.4.3. The homogenate was heated at 95^{o}C for 5 minutes. At this point an aliquot was taken and added in equal amounts to 2x Laemmli sample buffer and then heated again at 95°C for 5 minutes; this sample was then used for the tubulin loading control. The remaining sample was treated with 10µg/ml of proteinase K (Ambion) for 30 minutes at 37°C and then added in equal amounts to 2x Laemmli buffer. The sample was then heated at 95°C for 5 minutes and used to probe for DPR proteins. 2 µl of lysate was pipetted onto nitrocellulose membrane (Pall) and then left to dry at room temperature for 15 minutes. Membranes were then blocked in OBB mixed 1:1 with PBS for 1 hour at room temperature. Membranes were then left overnight at 4°C with primary antibody

diluted in OBB with 0.1% (v/v) tween-20 (Sigma-Aldrich). Following incubation with primary antibody membranes were washed 3x 5 minutes with PBS-T. Secondary antibodies were applied for approximately 1 hour at room temperature diluted to the required concentration in OBB with 0.1% (v/v) tween-20.

2.5.2 Antibodies used for dot blotting

The following primary antibodies and dilutions were used;

Antibody	Dilution	Source
Rabbit polyclonal anti-GR	1:1000	Kind gift from David Mann (University of Manchester)
Rabbit polyclonal anti-GR	1:1000	Kind gift from Leonard Petrucelli (Mayo Clinic)
Rabbit polyclonal anti-GP	1:5000	Eurogentec (Hirth/Gallo/Shaw labs, King's College London)
Mouse monoclonal anti-GP	1:1000	Abmart (Hirth lab, King's College London)
Rabbit polyclonal anti-PR	1:1000	Kind gift from Leonard Petrucelli (Mayo Clinic)
Rabbit polyclonal anti-PA	1:1000	Kind gift from Leonard Petrucelli (Mayo Clinic)
Mouse monoclonal anti-beta tubulin clone E7	1:300	Developmental Studies Hybridoma Bank

The following secondary antibodies and dilutions were used;

Antibody	Dilution	Source
Polyclonal goat anti-rabbit IgG (HandL) conjugated IRDye800	1:10000	Rockland immunochemicals, (611132002)
Polyclonal goat anti-mouse IgG (HandL) conjugated Alexa Fluor 680	1:10000	Thermo Fisher (A21057)

2.6 Immunohistochemistry (IHC) of *Drosophila* tissue

2.6.1 Solutions for IHC of Drosophila tissue

0.1M Sodium phosphate dibasic (Na_2HPO_4)

0.1M Sodium phosphate monobasic (NaH_2PO_4)

0.1M Sodium phosphate buffer (PB)

0.1M Na_2HPO_4 was added to NaH_2PO_4 to achieve a pH of 7.4

0.1M Lysine HCL (PBL) (100ml solution)

3.6% (w/v) Lysine HCL (Sigma-Aldrich)

Na_2HPO_4 was added to achieve a pH of 7.4

0.1M PBS added until volume was 100ml

2% and 4% Paraformaldehyde fix (PLP)

For larval brains, pupal eye discs and adult *Drosophila* brains:

75% (v/v) PBL

25% (v/v) 8% (w/v) paraformaldehyde (Sigma-Aldrich)

For *Drosophila* salivary glands:

50% (v/v) PBL

50% (v/v) 8% (w/v) paraformaldehyde (Sigma-Aldrich)

PB-Triton X (PBT)

0.5% (v/v) Triton X-100 (Sigma) in 0.1M PB.

PBT-normal goat serum (PBT-NGS)

10% (v/v) NGS (Gibco) in PBT

2.6.2 Larval brain, pupal eye discs and adult brain dissection and staining

Larval brains, pupal eye discs and adult brains were dissected in cold 0.1M PB in a time window of a maximum of 30 minutes. Following dissection, the tissue was fixed in 2% PLP and left rotating for 1 hour at room temperature. After fixation, the tissue was washed 3x 5min in PBT and then blocked in 10% PBT-NGS for 15 minutes. Samples were then agitated overnight rotating at 4^{o}C in the appropriate concentration of primary antibody diluted in 10% PBT-NGS. Proceeding overnight incubation, the primary antibody solution was removed by 3x 5 minute PBT washes and the secondary antibody added diluted in 10% PBT-NGS. Tissues were agitated for 3 hours in the dark at room temperature with secondary antibody, after which the secondary antibody was removed, and the tissue washed 2x 15 minutes in PBT followed by a final 2x 15 minutes wash in PB. Samples were then incubated overnight at 4^{o}C with Vectasheild containing 4',6-diamidino-2-phenylindole (DAPI) (Vector laboratories, H1200). Following incubation with DAPI the samples were then mounted on glass slides (Thermo Scientific) with cover slips placed on top sealed with clear nail varnish.

2.6.3 Salivary gland dissection and staining

Salivary glands were dissected and stained using a protocol adapted from Freibaum et al. (2015). Salivary glands were dissected in cold PB for total time of 15 minutes. Salivary glands were then fixed with 4% PLP for 15 minutes via rotation and then washed 3x 5 minutes in PBT. The samples were then blocked for 1 hour in 10% PBT-NGS and then left rotating overnight at 4^{o}C with primary antibody diluted in 10% PBT-NGS. After overnight primary antibody staining the salivary glands were washed 3x 10 minutes in PBT. Following PBT washes the tissues were left with secondary antibody for 1 hour agitating at room temperature. For a final step the secondary antibody was removed, and the samples washed 3x 10 minutes in PBS and then left at 4^{o}C overnight in Vectasheild w/ DAPI. The salivary glands were then mounted as described previously for other tissues.

2.6.4 Antibodies used for IHC of Drosophila tissue

Antibody	Dilution	Source
Rabbit polyclonal anti-GP	1:1000	Eurogentec (Hirth/Gallo/Shaw labs, King's College London)
Mouse monoclonal anti-GA clone 5E9	1:500	Merck Millipore (MABN889)
Rat monoclonal anti-GR clone 5A2	1:300	Merck Millipore (MABN778)
Mouse monoclonal anti-HA clone 6E2	1:500	Cell Signal (2367S)
Rabbit polyclonal anti-HA	1:500	Santa Cruz (sc-805)
Rabbit polyclonal anti-TBPH	1:2000	Hirth Lab (King's College London)
Rabbit polyclonal anti-RanGAP	1:300	Kind gift from C. Staber (Stowers Institute for Medical Research)
Rabbit polyclonal anti-Importin-α3	1:300	Kind gift from S. Cotterill (St George's)
Rabbit polyclonal anti-Pendulin	1:300	Kind gift from M. Frasch (Friedrich-Alexander University Erlangen-Nürnberg)
Mouse monoclonal anti-dRCC1	1:10	Kind gift from M. Frasch (Friedrich-Alexander University Erlangen-Nürnberg)
Mouse monoclonal anti-MAB414	1:500	Abcam (ab24609)
Rabbit polyclonal anti-dNup50	1:10000	Kind gift from J. Großhans (University of Göttingen)
Rabbit polyclonal anti-Ref(2)P	1:1000	Kind gift from D. Contamine (Versailles Saint-Quentin-en-Yvelines University)

The following secondary antibodies and dilutions were used;

Antibody	Dilution	Source
Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	1:150	Life Technologies (A-11034)
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	1:150	Life Technologies (A-11001)

Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 568	1:150	Life Technologies (A-11011)
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 568	1:150	Life Technologies (A-11031)
Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 647	1:150	Life Technologies (A-21245)
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 647	1:150	Life Technologies (A-21236)
Goat anti-Rat IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 647	1:150	Life Technologies (A-21247)

2.7 IHC staining of post-mortem human brain tissue

All staining was performed on 7µm frontal cortex tissue sections was provided by the MRC London Neurodegenerative Diseases Brain Bank, case details can be found in appendix 9; table A9.1.

2.7.1 Solutions for IHC staining of post-mortem human brain tissue

95% Industrial methylated spirit (IMS)

99% IMS (Solmedia) diluted in appropriate volume of dH_2O

70% Industrial methylated spirit (IMS)

99% IMS (Solmedia) diluted in appropriate volume of dH2O

Citrate Buffer pH 6.0

Sodium citrate tribasic dihydrate (Sigma) dissolved in dH2O

Adjusted to pH 6.0 with 1M HCL

Tris buffered saline (TBS) pH 7.4

Trizma base (Sigma)

Sodium Chloride (Sigma)

Phosphate buffered saline (PBS) pH 7.4

10x PBS tablets (Severn Biotech) dissolved in 1L ddH_20

3,3'-Diaminobenzidine (DAB) solution

1 DAB tablet (Sigma) dissolved in 20ml TBS

1% Acid Alchohol

70% IMS

200µl HCL (Sigma)

2.7.2 DAB immunolabeling

Frontal cortex sections were initially warmed on a hotplate. Sections were dewaxed by sequential washing in the following solutions. 2 minutes in xylene (Solmedia), followed by a 2 minute wash in fresh xylene. Tissue was then washed in in 99% IMS alcohol (Solmedia) for 2 minutes, another wash in fresh 99% IMS alcohol and finally a few seconds wash in 95% IMS. Hydrogen peroxide (VWR chemicals) was then used to quench the hydrogen peroxidases that naturally exist within tissue slices. Endogenous peroxidases were blocked by placing tissue sections in 300ml of ethanol (Solmedia) with 7ml of 30% aqueous hydrogen peroxide solution (VWR) for 30 minutes. Sections were then washed in running tap water and then washed in dH_2O. Citrate buffer treatment was then used to open antigen epitopes. Sections were placed in a plastic container filled with citrate buffer (pH 6.0) and the lid covered with Clingfilm. Sections were then microwaved at full power (800kW) for 8 minutes followed by being microwaved again 2x 8 minutes on simmer. Sections were cooled in cold running tap water and then washed 2x minutes in TBS, after which a hydrophobic wax ring was drawn around the tissue. The tissue was then blocked in the appropriate normal goat blocking serum (NGS) (Dako) diluted 1:10 in TBS for 20 minutes, after which the primary antibody was added at the appropriate

concentration diluted in TBS with NGS at 1:100. Sections were left overnight at 4ºC. Proceeding overnight incubation, the tissue was washed 2x 5 minutes in TBS and then left with the appropriate secondary antibody for 45 minutes at room temperature. The tissue was then washed 2 x minutes with TBS and then incubated with ABC HRP solution (ThermoFischer Scientific) prepared 40 minutes in advance. Sections were incubated with the ABC HRP solution for 45 minutes and then washed 2 x minutes in TBS. Filtered DAB solution activated with hydrogen peroxide was added to the tissue until appropriately stained. The excess DAB solution was then washed in cold running tap water for 10 minutes after which the sections were counter-stained with haematoxylin (Harris Haematoxylin) for 20 seconds – 1 minute. Following the haematoxylin counter-stain sections were washed in cold running tap water for 5 minutes and then placed in acid alcohol for 2 seconds, washed again in running tap water and then dH_2O. Sections were then dehydrated by washing sequentially in the following solutions for 2 minutes – IMS alcohol (70%) wash, IMS alcohol (95%) wash, 1st IMS alcohol (99%) wash, 2nd IMS alcohol (99%) wash, 1st Xylene wash, 2nd Xylene wash, 3rd Xylene wash and 4th Xylene wash. Tissue sections were then mounted using DPX mounting medium (VWR) and glass coverslips.

2.7.3 Double immunofluorescence staining of post-mortem human brain tissue

Frontal cortex sections were initially warmed on a hotplate. Sections were then dewaxed by being washed for 2 minutes in xylene, followed by another 2 minute wash in fresh xylene. After xylene washes the tissue was the further washed in 99% IMS for 2 minutes, proceeded by another wash in fresh 99% IMS and finally a 30 second wash in 95% IMS. IMS was removed with a 5 minute wash in dH_2O. Tissue sections then underwent citrate buffer treatment. Sections were placed in a plastic container filled with citrate buffer (pH 6.0) and the lid covered with Clingfilm. The sections were then microwaved at full power (800kW) for 8 minutes and then microwaved 2x 8 minutes on simmer. Tissue sections were then cooled down in cold tap water before being transferred to fresh dH_2O for 5 minutes. Following citrate buffer treatment sections were circled with hydrophobic wax and blocked for 45 minutes in normal goat serum (Dako) at room temperature 1:10 in PBS. Tissue

sections were then left overnight at 4^{o}C with primary antibodies diluted to appropriate concentration in 10% PBS-NGS. Following overnight incubation with primary antibody the tissue sections were washed 2 x 5 minutes in PBS. Secondary antibody was then applied also diluted to the required concentration in 10% PBS-NGS. Secondary antibodies were left on for 45 minutes at room temperature in the dark. Sections were then washed 3 x minutes in PBS and then underwent sudan black treatment. 0.15g of sudan black (Acros Organics) dissolved in 50ml 70% ethanol was prepared 45 minutes in advance. The sudan black solution was left to mix and was then filtered before being applied to the tissue sections. Sections were left with sudan black for 10 minutes before being washed 8 x 2 minutes with PBS. Tissue sections were then mounted with glass coverslips and Vectasheild containing DAPI.

2.7.3.1 Antibodies used for immunostaining of human tissue

The following primary antibodies and dilutions were used;

Antibody	Dilution	Source
Rabbit polyclonal anti-KPNA4	1:250	Novus Biologicals (NBP1-31260)
Mouse monoclonal phospho-TDP-43 (pS409/410-1)	1:500	Cosmo Bio Co., LTD (TIP-PTD-M01)
Mouse monoclonal anti-GP	1:1000	Abmart (Hirth lab, King's College London)
Mouse monoclonal anti-GA 5E9	1:500	Merck Millipore (MABN889)
Rat monoclonal anti-GR 5A2	1:50	Merck Millipore (MABN778)

The following secondary antibody and dilution was used for DAB immunolabeling;

Antibody	Dilution	Source
Biotinylated swine anti-rabbit immunoglobulin	1:100	Agilent Dako (P0399)

The following secondary antibodies and dilutions were used for double immunofluorescence;

Antibody	Dilution	Source
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	1:150	Life Technologies (A-11001)
Goat anti-Rat IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	1:150	Life Technologies (A-11006)
Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 568	1:150	Life Technologies (A-11011)

2.8 Fluorescence *in situ* hybridisation *(FISH)*

2.8.1 Solutions for FISH

RNase free PBS Phosphate-saline buffer (PBS) (ThermoScientific)

RNase free PBS Tween-20

RNase free PBS

0.1% (v/v) Tween-20 (Sigma)

RNase free PBS Triton X-100

RNase free PBS

0.5% (v/v) Triton-X-100

Prehybridisation buffer

2x Saline sodium citrate (SSC) (Sigma)

40% Formamide (Fluka)

Salmon sperm DNA (1mg/ml) (Ambion)

0.1% Tween-20

$(G_2C_4)_4$ Oligonucleotide probe (IDT)

Wash buffer 1

40% Formamide

2x SSC

0.01% Tween-20

Wash buffer 2

2x SSC

0.1% Tween-20

2.8.2 FISH of larval salivary glands

The protocol used adapted from Tran et al (2015). Salivary glands were dissected in cold PBS and fixed in 4% PFA for 20 minutes. Samples were washed three times for 10 minutes in RNase-free PBS containing 0.1% Tween-20 and then permeabilised with PBS-0.5% Triton X-100 for 10 minutes followed by a 30-minute incubation in prehybridisation buffer at 55°C. Samples were incubated for 2 hours at 55ºC in prehybridisation buffer containing an Alexa488-labelled $(G_2C_4)_4$ RNA probe (0.2ng/µl). Samples were then washed in pre-warmed wash buffer A for 30 minutes at 55ºC and twice in wash buffer B for 30 minutes at room temperature. After washing with 0.1% Tween-20 once for 10 minutes, samples were blocked and probed as described for IHC (section 2.6.3) for mouse anti-phospho-RNA polymerase II 4H8 (1:500, Abcam) and secondary antibody used was Alexa fluorochrome 568 (Life Technologies). FISH stained salivary glands were mounted on glass slides in Vectasheild w/ DAPI as described previously.

2.9 Semithin retinal sections

2.9.1 Solutions for semithin retinal sections

Fixative solution

2% glutaraldehyde (Fluka) in 0.1M PB pH 7.2

Post-fixative solution

50% (v/v) osmium (Electron Microscopy Sciences),
50% (v/v) PB (0.2M)

Resin

Resin is composed of four components:

Epoxy resin (A = 54g), hardener (B = 44.5g), accelerator (C = 2.5g) and plasticiser (D = 10g) (Durcapan ACM, Sigma Life Science).

1% Toluidine-blue in borax

1g Toluidine blue
2g Borax (Sigma)
Made up to 100ml with deionised water

2.9.2 Semithin retinal sections

Flies were anaesthetised with CO_2 and then decapitated with a scalpel. Heads were then split in two with the right eye being kept for sections. 35 days old males' eyes were incubated in 100µl of the fixative solution and then centrifuged at 14,000rpm for 30 seconds at 4°C. The same volume of post-fixative solution was added for 30 minutes on ice; this solution was then replaced with 100% of post-fixation solution for 5 hours. Progressive dehydration by ethanol (30%, 50%, 70%, 90%, and 100%) was proceeded prior to two 10 minute propylene oxide washes at room temperature. Eyes were then incubated in mix (1:1) of propylene oxide (Sigma-Aldrich) and epoxy resin overnight. Retinas were embedded the following day in 100% epoxy resin between 4 and 6 hours at room temperature and thereafter baked at 80°C overnight.

The blocks were then trimmed to remove any excess resin around the samples and the surfaces smoothened to help visualization under the ultramicrotome. Sections of 1µm sections were performed using a ultramicrotome (Leica Ultracut UTC). Floating sections were placed onto a glass slide and then dried on a hot plate and stained with 1% toluidine-blue in borax using a filtered syringe to apply the solution onto the samples. The slides were then put again onto the hot plate at 100°C for 10 seconds followed by washing with deionised water. After the slides were dried they were mounted in DPX (Sigma), covered with coverslips and left to harden. Images were taken with light microscopy (AxioImager Z1, Zeiss). Quantification has been performed by counting the number of photoreceptors on 100 ommatidia on average per picture using ImageJ.

2.10 Confocal microscopy and image processing

For all immunofluorescent staining of *Drosophila* tissue and human brain images were obtained using a Lecia TCS SP5 confocal microscope with Leica Application Suite Advanced Florescence (LAS-AF) software (version 2.0.2). The fluorescent channels were scanned sequentially to prevent bleed-through. All tissues were scanned at a resolution of 1024x1024 pixels. Images and Z-projections were processed using ImageJ (http://imagej.nih.gov/ij/).

2.11 Behavioural assays

2.11.1 Startle-Induced Negative Geotaxis

The startle-induced negative geotaxis protocol is used to assess the climbing performance of flies. Startled-induced negative geotaxis was performed using modified 25ml serological pipettes (Costar). The bottom of the pipette was sealed with parafilm (Alpha Laboratories) and the top cut off and sealed with cotton wool. A maximum of 15 flies (n_{total}) were anaesthetised and gently placed into the pipettes and then left to recover and acclimatise for approximately 30 minutes in a 25°C incubator. For an initial test trial flies were firmly tapped to the bottom and then allowed to climb for 45 seconds. This was the repeated a further 3 times; however, for these test trials after 45 seconds the total number of flies below the 2ml line

(n_{bottom}) and the total number of flies that were above the 25ml line (n_{top}) was counted. Furthermore a 90 second recovery period was given between each trial. Based on the n_{top} and n_{bottom} scores a performance index was calculated for the three trials using the following equation:

$$PI = 0.5 \times (n_{total} + n_{top} - n_{bottom}) / n_{total}$$

All geotaxis was performed at the same 4 hour interval each day (between 12:00pm – 4:00pm). Only mated females were tested.

2.11.2 Video Assisted Motion tracking

All motion tracking was performed on mated females only. Flies were anesthetized and placed in tracking arenas; modified 6 well plates (diameter of 35mm) filled with silicone rubber (Sylguard) leaving approximately 3mm of space for flies to walk but not jump or fly. Flies were anaesthetised briefly and then transferred into the wells of the tracking arena. Flies were left to recover and acclimatise in a temperature controlled incubator (Stuart Scientific) set to 25°C for 45 minutes. All tracking was performed at 25°C in this temperature controlled incubator. After acclimatisation fly movement was recorded for approximately 30 minutes using a monochrome camera (Point Grey, FL3-FW-14S3M-C) mounted above the tracking arenas. Files were recorded using the FlyCap2 (http://flycap2.software.informer.com) and Fview (http://code.astraw.com/projects/motmot/fview.html) software. All recordings were performed during the first hour of the light/dark cycle dark phase. Fly movie format (.fmf) videos of fly walking behaviour were loaded onto Ctrax software (http://ctrax.sourceforge.net) which allows analysis of the position of the flies during the video recordings. This positional data was the exported and analysed in MATLAB (Mathworks, version 2014a). Any errors in the tracking of the flies was fixed with the FixErrors GUI (Branson et al., 2009**,** http://ctrax.sourceforge.net/fixerrors.html). Fixed trajectories were then analysed in MATLAB using custom script written by Dr. Dickon Humphrey and Jonah Dearlove (Hirth lab). Such analysis provided mean activity, activity vs inactivity and activity over time for individual flies recorded. Activity was defined as movement per frame above a velocity of 2mm/second. Percentage of time active, activity per minute and

raster plots for each fly were generated in MATLAB from custom code. Statistical analysis was performed by exporting data to GraphPad Prism 6.

2.12 Statistics

GraphPad Prism 6 was used for the statistical analyses performed within the results section. For a comparison of means with less than 2 experimental conditions an unpaired t-test was used. A chi-square test was used for the semithin retinal sections to determine whether difference between the expected number of photoreceptors and the observed number of photoreceptors differed. For a comparison of means with numerous experimental conditions (n>2) and one independent variable, a one-way analysis of variance (ANOVA) was utilised. For a comparison of means with numerous experimental conditions (n>2) and two independent variables, a two-way ANOVA was utilised. Post-hoc analysis was performed using Fisher's least significant difference (LSD). Multiple comparisons were corrected for using the Bonferroni-Holm method, in which non-adjusted p-values are ranked in order, from lowest to highest. By adjusting for the number of experimental conditions the alpha level is corrected and p-values are tested agaisnt alpha in a stepwise manner. The null hypothesis is rejected if p > alpha for a pairwise comparison.

Chapter 3: Characterisation of RAN translated dipeptide repeat proteins in a *Drosophila* models of the G_4C_2 repeat expansion

Three non-mutually exclusive mechanisms of toxicity have been attributed to the G_4C_2 repeat expansion; a loss of *C9ORF72* protein function, G_4C_2 RNA toxicity and DPR toxicity. *Drosophila* have no homologue of *C9ORF72*, hence fly models of the G_4C_2 expansion have focused on the two pathogenic gain of function mechanisms resulting from the expression of repetitive G_4C_2 RNA. The overwhelming evidence from *Drosophila* models of this hexanucleotide repeat expansion points toward the DPRs being the major drivers of pathogenesis in ALS and FTD. This chapter characterised the DPRs produced in a novel *Drosophila* model of flies expressing different lengths of G_4C_2 RNA.

3.1 G_4C_2 repeat containing constructs used in this study

Multiple transgenic flies containing different lengths of the G_4C_2 repeat sequence were used in this study. In these flies the repeat sequence is located in the 3' untranslated region (3'UTR) of the red fluorescent protein DsRed2 (figure 3.1). These constructs were designed based on a similar approach used by Li et al. (2008) and Yu et al. (2011) to model CAG and CUG RNA expansions in *Drosophila* for spinocerebellar ataxia type 3 (SCA3) and Myotonic Dystrophy Type 1 (DM1) respectively. In both these studies the repetitive RNA was cloned into the 3'UTR of DsRed2, a protein with no connection to disease pathogenesis and was shown not to confound repeat associated toxicity. Plasmids containing 8 (denoted as 8B), 38 and 56 G_4C_2 repeats in the 3'UTR of EGFP were generated by Younbok Lee (Shaw Lab) and published in Lee et al. (2013). The EGFP was removed and the repeats inserted into the 3'UTR of DsRed2 contained within the pUAST *attb* vector for site specific integration into the *Drosophila* genome by Yoshitsugu Adachi (Hirth lab). The 8 (denoted as 8A), 32, 64 and 128 DsRed2 repeat containing constructs were generated in the Hirth lab by Alan Stepto and Yoshitsugu Adachi via sequential doubling of 4x G_4C_2 repeats ligated into the 3'UTR of *DsRed2* in the pUAST *attb* vector. Confirmation of insertion into the ZH86FB locus was confirmed by Alan Stepto via RT-PCR. Repeat size was confirmed by Alan Stepto and Brad Smith (Shaw Lab) via Sanger sequencing and repeat primed PCR respectively. Expression of *DsRed2* RNA containing G_4C_2 repeats was confirmed via both southern and northern blotting by Alan Stepto. Due to repeat instability, the 128 repeat line was excluded from most experiments.

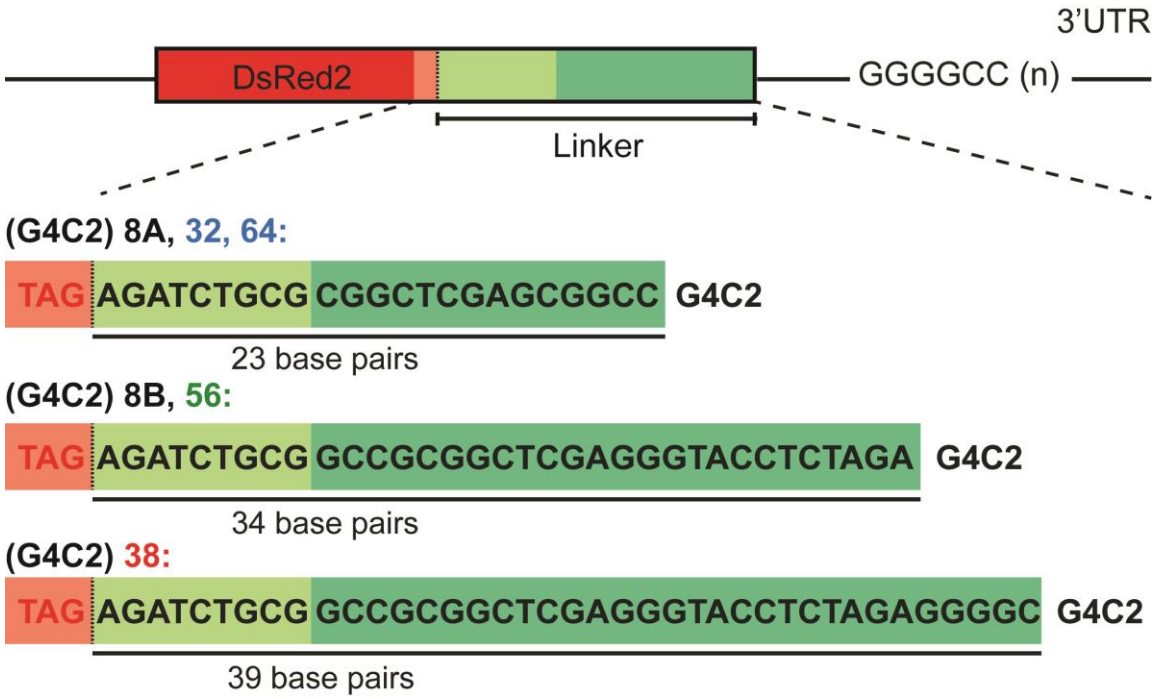

Figure 3.1 | Schematic of the G_4C_2 constructs used in this study. Schematic of generated constructs with different lengths of uninterrupted G_4C_2 repeats cloned into 3' UTR of the disease-unrelated marker gene *DsRed2*; variable linker sequences between *DsRed2* stop codon and start of G4C2 repeats are indicated for each construct.

3.2 G_4C_2 RNA foci cannot be detected in G_4C_2 repeat expressing larval salivary glands

One potential pathogenic mechanism derived from the G_4C_2 expansion is the formation of nuclear G_4C_2 RNA foci and the subsequent sequestration of RNA binding proteins into these foci. As later experiments in this thesis utilise salivary gland cells of L3 *Drosophila* larvae to investigate DPR toxicity; RNA foci were probed for in this tissue using the *FKH-GAL4* driver. Foci were investigated in salivary glands in all repeat containing flies and a positive control - a G_4C_2 RNA-only fly line previously characterised and validated in Mizielinska et al. (2014). This RNA-only construct consists of a 288 G_4C_2 expansion with base pair interruptions containing stop codons inserted every 12 G_4C_2 repeats, leading to a stop codon in all sense and antisense reading frames. This construct still forms stable RNA G-quadruplexes and RNA foci but does not produce detectable levels of DPRs (Mizielinska et al, 2014). Fluorescence *in situ* hybridisation (FISH) using an Alexa488-labelled $(G_2C_4)_4$ RNA probe was performed. No signal was observed in negative controls (*FKH/+* and *FKH>DsRed2*) or in salivary glands cells with 8 G_4C_2 repeats (*FKH>(G_4C_2)8B*) (figure 3.2). The 288-RNA only control showed multiple large and small foci in a single salivary gland cell (figure 3.2). FISH in salivary glands expressing 32 or more repeats showed a single large foci per cell (figure 3.2). Both Mizielinska et al. (2014) and Tran et al. (2015) previously also found a single large foci in salivary gland cells expressing G_4C_2 repeats. Tran et al. (2015) found in flies with intronic 160 G_4C_2 repeats flanked by parts of the *C9ORF72* human intronic and exonic sequence that these single foci co-localised with activated Pol-II; indicating they represent the site of transgene transcription rather than a “true” RNA foci. To explore this possibility in the repeat flies used in this study FISH coupled with fluorescent immunohistochemistry was performed, where the repeat containing salivary glands were co-stained with the Alexa488-labelled $(G_2C_4)_4$ RNA probe and an antibody for phosphor RNA polymerase II (p-Pol-II). Co-localisation between the single large foci and p-Pol-II as was reported by Tran et al., (2015) was seen in salivary gland cells expressing 32 or more repeats. Therefore, the signal detected by $(G_2C_4)_4$ RNA probe in the repeat flies may represent the start site of G_4C_2 RNA transcription rather than a true RNA foci.

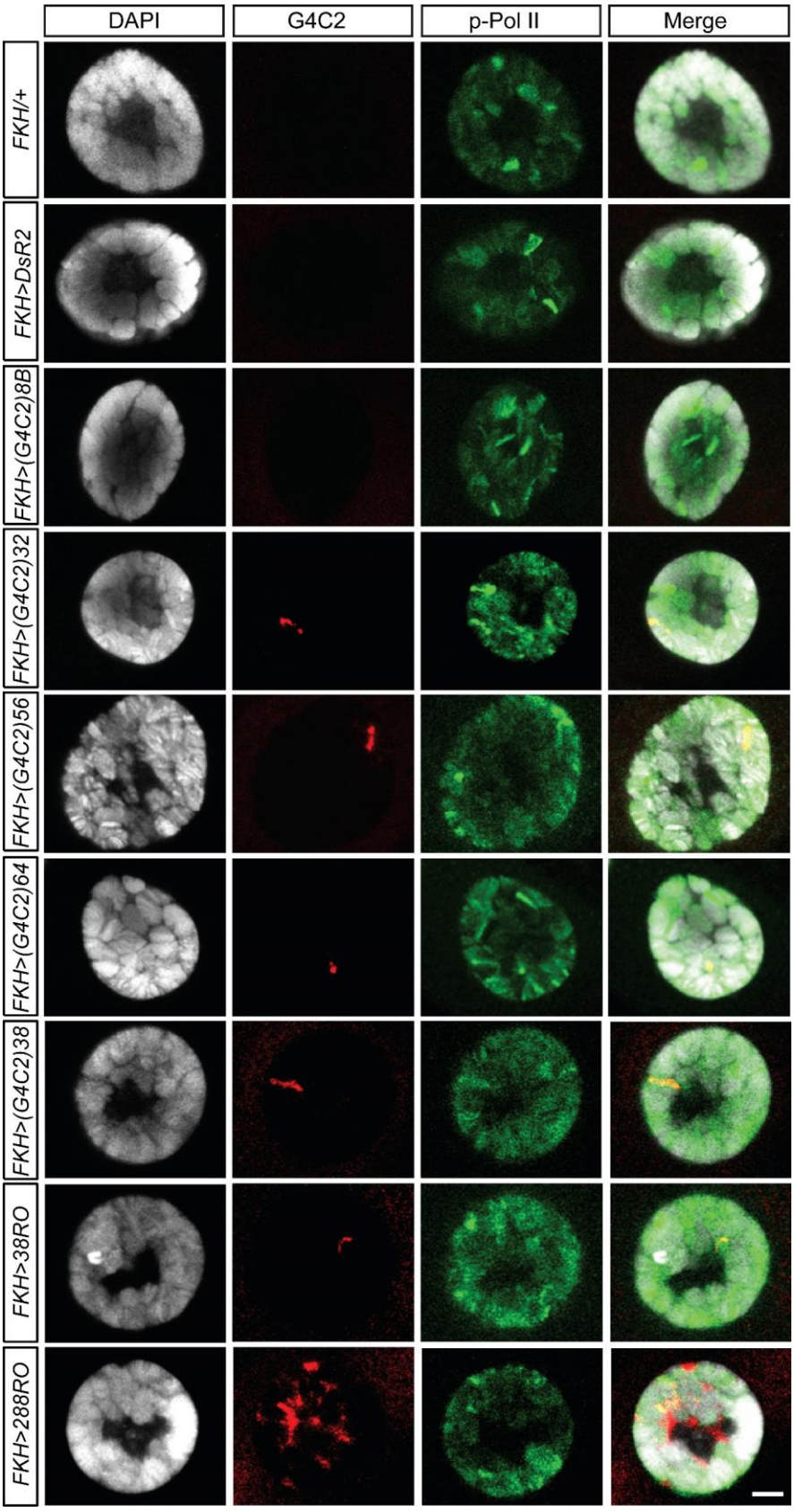

Figure 3.2 | True sense RNA foci are not detected in larval salivary glands of pure repeat flies. Fluorescence *in situ* hybridization of a $(G_2C_4)_4$ RNA probe together with immunostaining for phospho RNA polymerase II (p-Pol-II) identifies co-localisation at site of transcription in cells expressing 288 G4C2 repeat RNA only, as well as for 32, 56, 64, 38 repeats and 36 G_4C_2 repeat RNA only. Additional RNA foci non-overlapping with p-Pol-II is seen only for 288RO. Scale bar, 10µm.

3.3 No read-through translation of DsRed2 stop codon in G_4C_2

The location of the repeats in the 3'UTR of *DsRed2*, and the presence of a TAG stop codon before the 3'UTR is expected to prevent translational read-through of the repeat sequence. Hence to confirm no DsRed2-DPR fusion proteins were translated in repeat containing flies, western blots were performed on lysate generated from the heads of flies expressing the differing DsRed2 constructs in post-mitotic neurons using the *Elav^{C155}*-GAL4 driver (figure 3.3). Membranes were probed with antibody specific for DsRed2 with anti-beta Tubulin as a loading control. DsRed2 as expected was not seen in the negative control genotype (*Elav^{C155}*/+) and runs at its predicted molecular weight of 27.6kDa in the DsRed2 positive control (*Elav^{C155}>DsRed2*). The presence of an identical weight band in all DsRed2-G_4C_2 repeat containing fly lines no matter the length of repeat indicates little or no translational read-through is occurring (figure 3.3). DsRed2 does not run higher than its molecular weight at 27.6kDa. Hence any DPRs detected in flies harbouring G_4C_2 repeats are likely a product of RAN translation rather than read-through translation of the repeats.

DsRed2 protein levels were significantly increased in all repeat containing lines compared to the DsRed2 control (appendix 1; figure A1; quantification performed by Alan Stepto). The levels of DsRed2 also differed between different repeat lines (appendix 1; figure A1). DsRed2 protein levels were significantly higher in the 38 repeat flies compared to the 8A, 8B, 32, 56 and 64 repeat lines. There was no significant difference in DsRed2 protein levels between the 32, 56 and 64 repeat lines. The DsRed2 levels were also significantly higher in the 32 and 64 repeat lines compared to the 8A and 8B lines but this was not the case for the 56 repeat line.

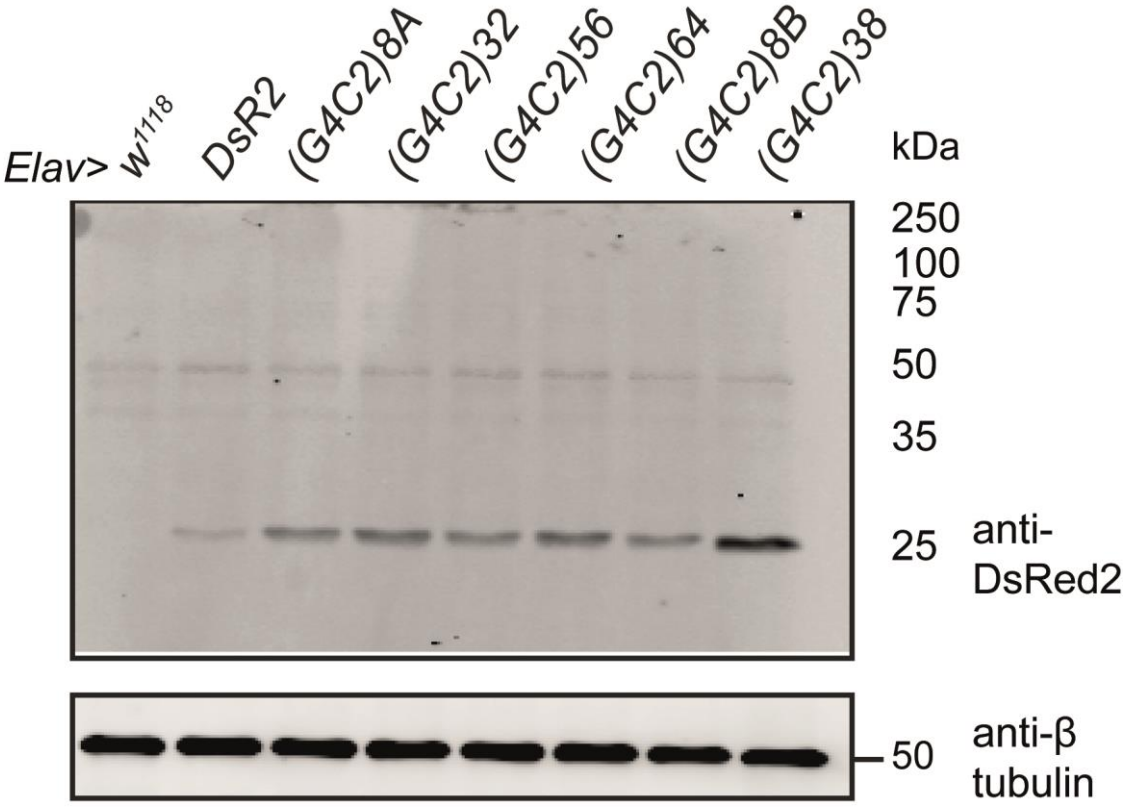

Figure 3.3 | No higher molecular weight DsRed2 protein is detected in flies harbouring G_4C_2 repeats. Western blot analysis of RIPA soluble protein from day 5 head extracts showing the expected ~26 kDa DsRed2 band in all G_4C_2 construct expressing flies, but not in negative control (*Elav*^{c155}/+), indicating no read-through translation of the *DsRed2* stop codon into the G_4C_2 repeats.

3.4 Poly-GP is translated in a length and construct dependent manner

To assess whether the G_4C_2 mRNA was subject to RAN translation in this model, the pan-neuronal driver *Elav^{C155}-Gal4* was used, and protein lysate from 5-day old fly heads was probed with a novel monoclonal mouse poly-GP antibody (the characterisation of these antibody can be found in appendix 3, (figure A3 A and A3 B). Poly-GP not detected in controls (*Elav^{C155}*/+ and *Elav^{C155}>DsRed2)* or in flies harbouring ≤8 repeats, however was detected at variable levels in all flies harbouring ≥32 repeats. This is consistent with the hypothesis that there is a length-dependent threshold for RAN translation initiation, and that this threshold may lie between 8 and 32 G_4C_2 repeats.

Interestingly significantly higher levels of poly-GP were observed in the 32 and 64 G_4C_2 repeat lines, with much lower levels seen in the 38 and 56 G_4C_2 repeat lines (figure 3.5B). Hence the expression of poly-GP detected in each condition varied dramatically in both a length and construct dependent manner. Furthermore, the poly-GP bands in the 32 (~37 kDa and ~ 30 kDa), 56 (~50 kDa), 64 (many bands detected between ~30-50 kDa) and 38 (~38 kDa) repeat lines were detected at a much higher molecular weight than expected assuming translation initiates at the first G_4C_2 repeat and ends at the first stop codon in the 3' UTR (13.87 kDa; 24.29 kDa; 27.77 kDa; and 16.47 kDa respectively). Quantification of poly-GP western blots were performed in collaboration with Alan Stepto (figure 3.5B).

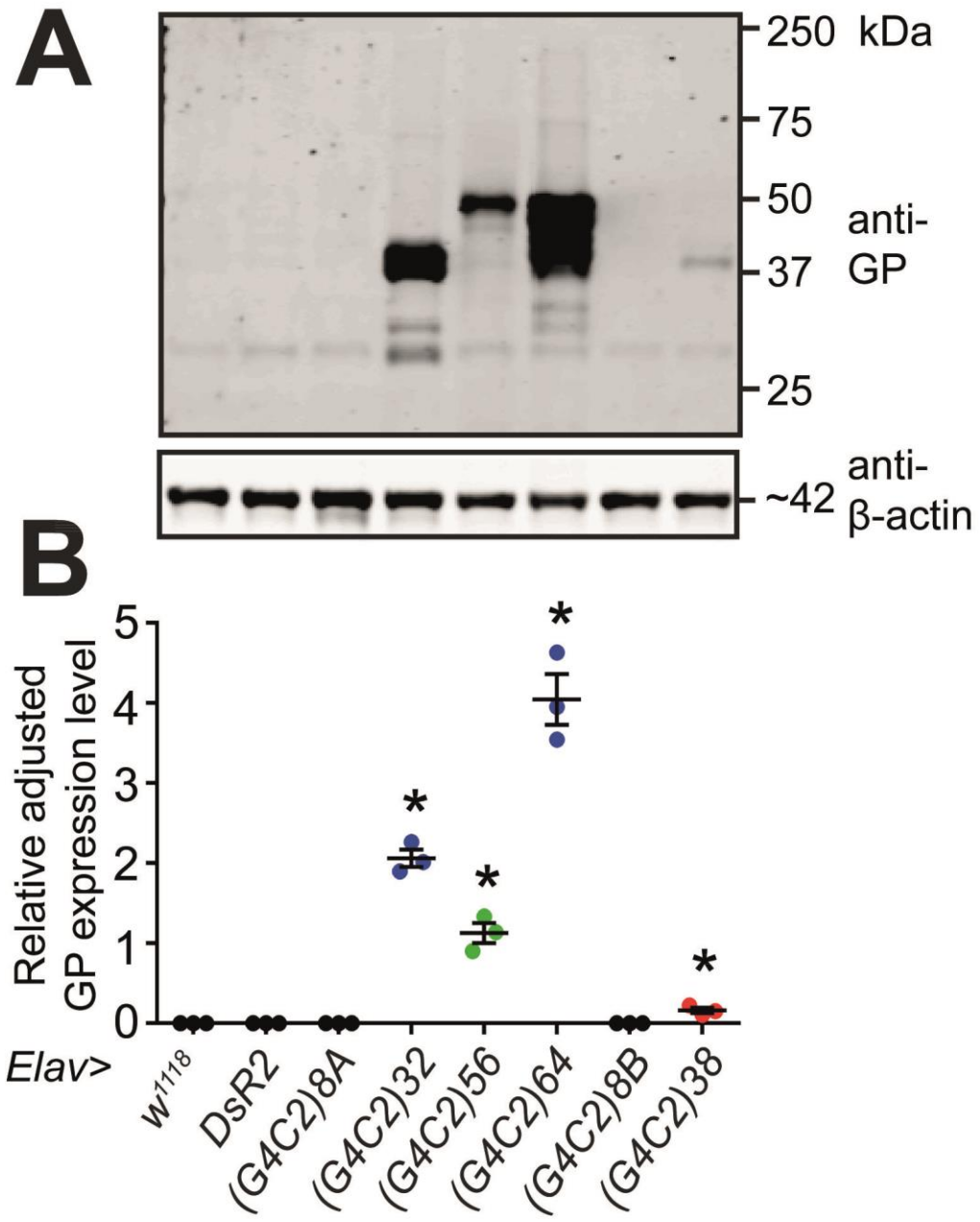

Figure 3.5 | Translation of G4C2 derived poly-GP. **(A)** Western blotting and quantification of poly-GP expression in 5-day old flies expressing G_4C_2 repeats pan-neuronally ($Elav^{C155}$-Gal4). Poly-GP was detectable in the 32, 56, 64 and 38 G_4C_2 repeat lines at differing sizes. No poly-GP signal was detected in the negative controls (*Elav* ^{C155}/+ and *Elav* ^{C155}*>DsR2*) or in flies expressing 8 repeats (*Elav* ^{C155}*>(*G_4C_2*)8A*). **(B)** Quantification of western blots for expression of poly-GP. A one-way ANOVA demonstrated there to be a significant difference between the genotypes for poly-GP expression levels (F=133.5, P<0.0001). Bonferroni-Holm corrected multiple comparisons can be found in appendix 7, table A7; (*denotes significant difference from other genotypes). Mean with SEM is shown for each genotype (n=3).

3.5 Poly-GA is translated in a length and construct dependent manner

The two most prevalent DPRs produced from the G_4C_2 expansion is poly-GA from the sense strand (Schludi et al., 2015). Hence whether the repeat flies used in this study also produce poly-GA was investigated using the mouse monoclonal anti-GA antibody (5E9) generated and validated by Mackenzie et al. (2013). Western blotting of protein lysate from 5 day old *Drosophila* heads expressing the repeats specifically in neurons using the *Elav^{C155}* driver revealed RAN translation of poly-GA in a length and construct dependent manner as was observed for poly-GP (figure 3.6A). No Poly-GA was detectable in the two negative controls (*Elav^{C155}*/+ and *Elav^{C155}>DsRed2*) or in flies with 8 repeats. Poly-GA was detectable in the 32 (~35 kDa), 38 (~37 and ~40 kDa), 56 (~40 kDa) and 64 repeat (~37 and ~45 kDa) containing lines. The levels of poly-GA expression followed the same pattern as was seen for poly-GP. Poly-GA protein levels were significantly higher in the 32 and 64 repeat flies compared to both 56 and 38 repeat flies (figure 3.5B). As was the case for poly-GP, poly-GA ran at a higher molecular weight than predicted for the 32, 56, 64 and 38 repeat constructs (13.92 kDa; 24.35 kDa; 27.83 kDa and 16.53 kDa respectively). As is the case for poly-GP, multiple band sizes of poly-GA are seen, specifically in the 64 and 38 repeat lines. Quantification of poly-GA western blots were performed in collaboration with Alan Stepto (figure 3.5B).

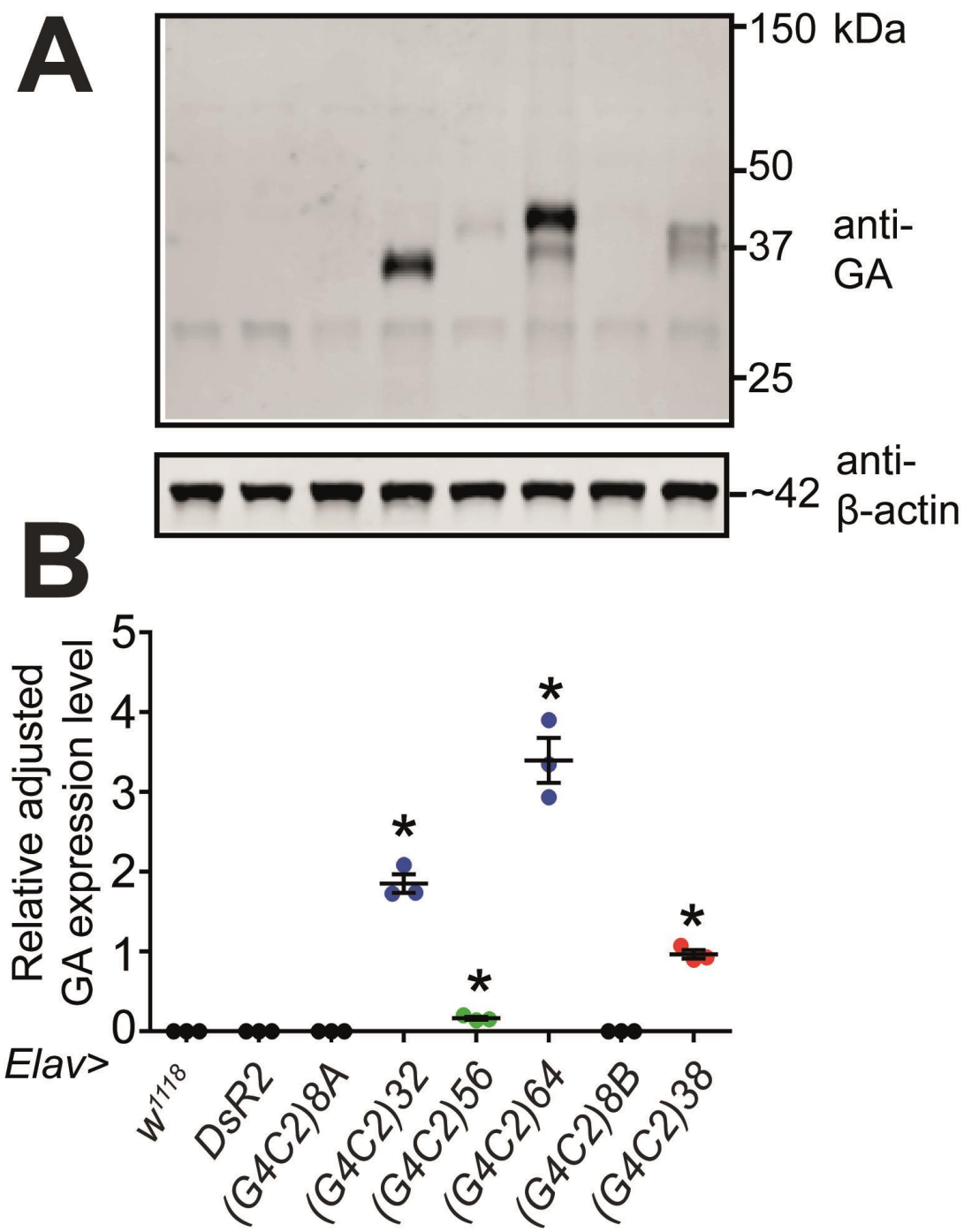

Figure 3.6 | Translation of G4C2 derived poly-GA. (A) Western blotting and quantification of poly-GA expression in 5-day old flies expressing G_4C_2 repeats pan-neuronally (Elav^{C155}-Gal4). Poly-GA was detectable in the 32, 56, 64 and 38 G_4C_2 repeat lines at differing sizes. No poly-GA signal was detected in the negative controls (*Elav ^{C155}/+* and *Elav ^{C155}>DsR2*) or in flies expressing 8 repeats (*Elav ^{C155}>(G_4C_2)8A*). **(B)** Quantification of western blots for expression of poly-GA. A one-way ANOVA demonstrated there to be a significant difference between the genotypes for poly-GA expression levels (F=80.04, P<0.0001). Bonferroni-Holm corrected multiple comparisons can be found in appendix 7, table A7; (*denotes significant difference from other genotypes). Mean with SEM is shown for each genotype (n=3).

3.6 Subcellular localisation of poly-GP and poly-GA differs between G_4C_2 repeat containing fly lines

In the human brain, poly-GP and poly-GA form neuronal cytoplasmic inclusions, neuronal intranuclear inclusions and dystrophic neurites particularly in the frontal cortex and cerebellum (Schludi et al., 2015). Immunofluorescent staining of *Drosophila* L3 larval brains was performed to determine the subcellular localisation of poly-GP and poly-GA *in vivo* using the rabbit poly-GP antibody (Hirth/Gallo/Shaw labs) and the poly-GA antibody generated by Mackenzie et al. (2013). No poly-GP or poly-GA was detected in control larval ventral nerve cord (VNC) neurons that did not express the hexanucleotide sequence (*Elav ^{C155}>DsRed2*) (figure 3.7). DPRs were however detected in the larval VNC of the 56 and 64 G_4C_2 repeat lines. Interestingly, the poly-GP and poly-GA observed in VNC neurons of larvae expressing 56 and 64 repeats displayed differential sub-cellular localisations (figure 3.7). For the 56 repeat larval VNC, poly-GP was diffusely nuclear, with regions of intense nuclear staining also observed. Poly-GP in the 64 repeat line was predominantly diffusely cytoplasmic with less intense diffuse nuclear staining. No regions of intense nuclear staining were observed in the 64 repeat VNC neurons. Thus, poly-GP displayed a completely different pattern of localisation between the 56 and 64 repeat conditions. The relatively low levels of poly-GA in the 56 repeat line compared to the 64 repeat line, as evidenced by SDS-PAGE, was not detectable by immunofluorescent staining. At such low levels a neuron may be able to efficiently clear the peptide preventing it accumulating via protein degradation pathways. In flies expressing 64 repeats however, poly-GA had a diffuse nuclear staining with compact regions of intense staining in nuclei. No cytoplasmic poly-GA was detected in flies expressing 64 repeats.

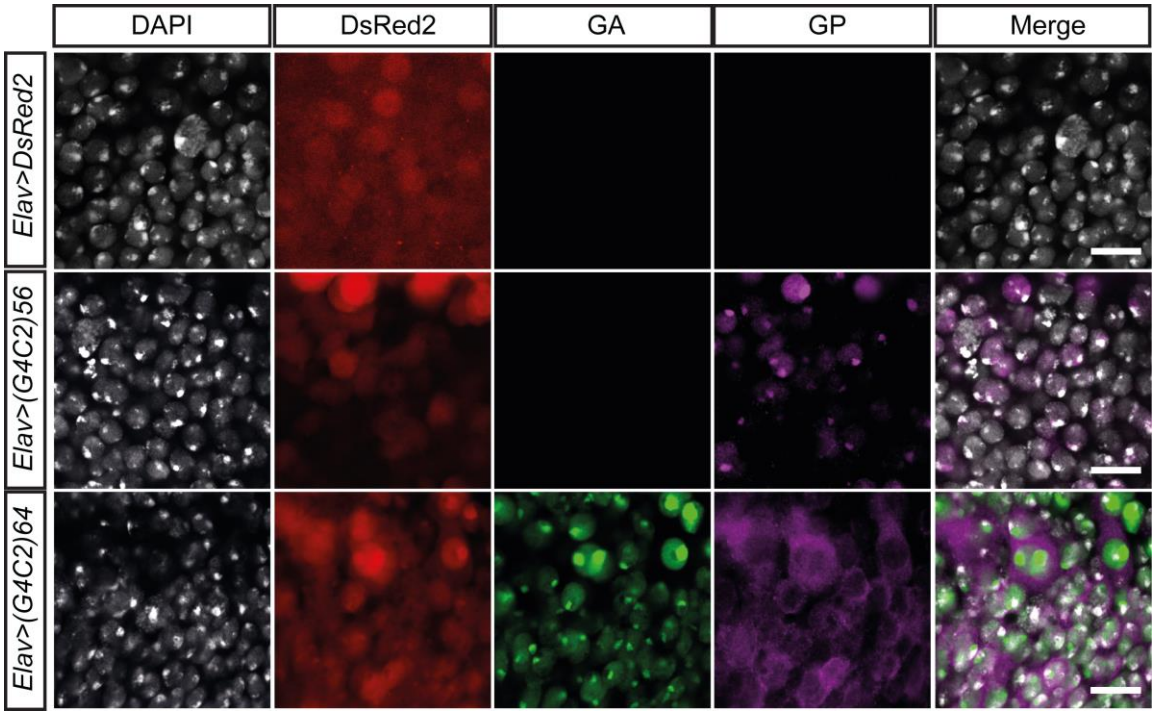

Figure 3.7 | Subcellular localisation of poly-GP and poly-GA in the L3 larval ventral nerve cord. Poly-GP and poly-GA are not detected in the DsRed2 negative control larval ventral nerve cord (VNC) (*Elav ^{C155}>DsRed2*). Poly-GP has a diffuse nuclear localisation in addition to nuclear accumulations in the 56 repeat larval VNC (*Elav^{C155}>(G_4C_2)56*). Poly-GA was not detectable by immunohistochemistry in the 56 repeat VNC. Poly-GA in the 64 repeat larval VNC (*Elav ^{C155}>(G_4C_2)64*) was diffusely nuclear with nuclear accumulations; poly-GP had a predominantly cytoplasmic stain with also some weak diffuse nuclear labelling also apparent. Scale bars, 10μm.

3.7 Poly-GR is only detected in 38 G_4C_2 repeat fly line

The third sense DPR translated from the G_4C_2 expansion is the arginine rich DPR poly-GR. Poly-GR inclusions are observed much less frequently compared to poly-GP and poly-GA in post mortem material (Schludi et al., 2015., Mackenzie et al., 2015). However, more recently it has been demonstrated that of the five DPRs, only poly-GR is significantly found more often in clinically relevant brain areas than in unrelated brain areas of ALS patietns (Saberi et al., 2018), furthermore poly-GR was shown to form dendritic-like aggregates within the motor cortex which co-localized with pTDP-43 inclusions (Saberi et al., 2018). This data combined with the disease models showing toxicity of the arginine DPRs (Mizielinska et al., 2014) makes the detection of poly-GR in the flies used in this study important due to its increasing association with both disease anatomy and toxicity. Owing to difficulties in detecting poly-GR by SDS-PAGE, due to the highly positive charge of the dipeptide, a dot blot protocol previously developed by Mizielinska et al., (2014) to detect RAN translated poly-GR in flies was performed. Dot blots were performed using two previously published poly-GR antibodies (Gendron et al., 2013; Mann et al., 2013) (figure 3.8). Lysate from HEK293 cells (provided by Younbok Lee) was used as a positive control to confirm the specificity of the poly-GR antibodies. Positive control lysate confirmed the poly-GR binding capacity of the antibodies used to probe fly lysate. No poly-GR was detectable in the negative controls (*Elav^{C155}*/+ and *Elav^{C155}>DsRed2*). No poly-GR signal above background seen in controls was detected for the 8B, 32, 56, 64 and 128 repeat lines. Poly-GR signal above background was however seen in the 38 G_4C_2 repeat fly line for both poly-GR antibodies tested; suggesting that only the 38 repeat line produced poly-GR.

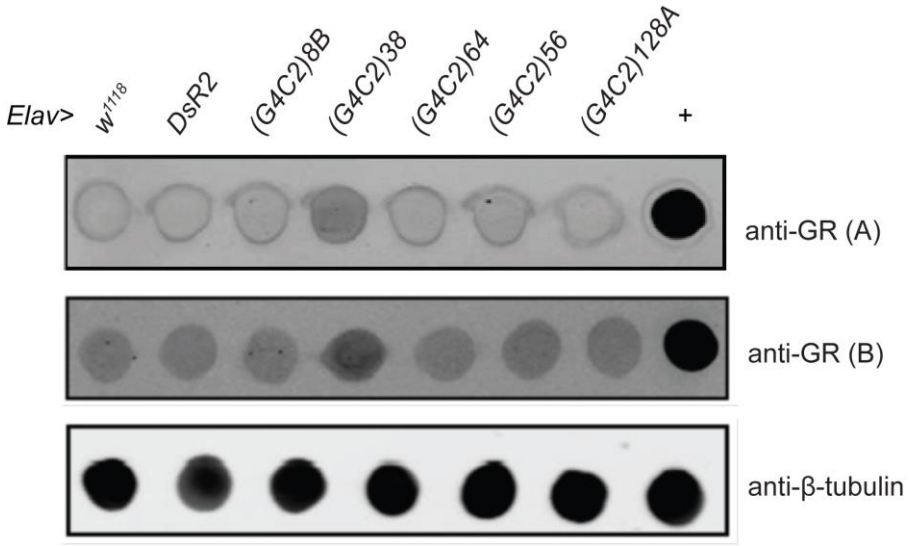

Figure 3.8 | Translation of G4C2 derived poly-GR. Two published poly-GR antibodies **(A)** Gendron et al., 2013) and **(B)** Mann et al., 2013) were used to detect the presence of poly-GR in RIPA soluble lysate from flies expressing 38 G4C2 repeats but not other repeat lines. No poly-GR was detectable in negative controls (*Elav* ^{C155}/+ and *Elav* ^{C155}*>DsR2*). Positive control of lysate from HEK293 cells expressing poly-GR from a non G_4C_2 template was used. Samples were also probed with anti-β-tubulin to confirm approximately equal loading of protein.

3.8 Subcellular localisation of poly-GR in 38 repeat flies

A different poly-GR antibody, rat monoclonal 5H9 (Mori et al., 2013) has been determined to be the most sensitive antibody for detecting poly-GR inclusion bodies in patient tissue compared to other available poly-GR antibodies (Mackenzie et al., 2015). Rat monoclonal 5H9 was used to probe day 5 adult fly brains for poly-GR (figure 3.9). Dissections for poly-GR localisation were performed in collaboration with Alan Stepto. No poly-GR signal was detectable in the negative control (*Elav* ^{C155}*>DsR2*) or the 8A, 8B, 32, 56 and 64 fly lines. 5H9 immunohistochemistry revealed a diffuse perinuclear staining pattern in the 38 repeat flies. Small accumulations of poly-GR were also seen, mainly perinuclear, but also occasionally in the nucleus itself.

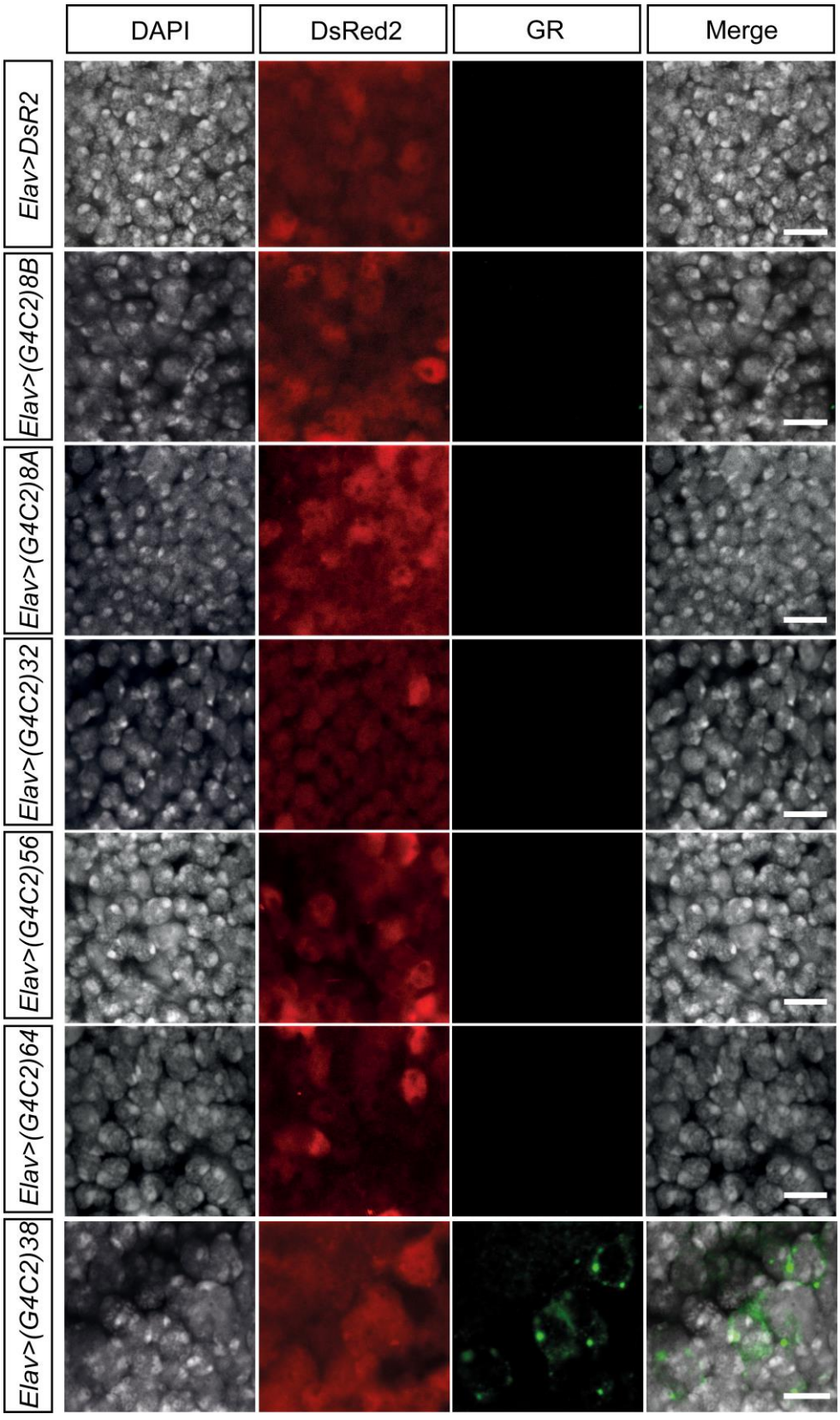

Figure 3.9 | Poly-GR is only detectable in the 38 repeat fly brain. Immunohistochemistry from 5-day old adult brains shows poly-GR is only detected in flies expressing 38 G4C2 repeats but is not detected in the negative control (*Elav* ^{C155}*>DsR2*) and flies expressing 8A. 8B, 32, 56 or 64 G4C2 repeats. Poly-GR displays a diffuse perinuclear localisation with regions of intense staining. Scale bars, 10µm.

3.9 Anti-sense DPRs are not detected in G_4C_2 fly lines

In addition to translation from the sense G_4C_2 strand, the anti-sense C_4G_2 is transcribed and translated to produce three anti-sense DPRs – poly-GP, poly-PR and poly-PA. Anti-sense RAN translation was probed in the repeat flies via dot blotting using previously validated and published poly-PR and poly-PA used to detect protein deposits in human post mortem tissue (Gendron et al., 2013). Day 5 lysate from fly heads expressing the repeats in neurons were probed for poly-PR and poly-PA (figure 3.10). No poly-PR or poly-PA signal was detected in controls (*Elav^{C155}*/+ and *Elav^{C155}>DsRed2*) or the 8A, 8B, 32, 56, 64 and 38 repeat fly lines. Hence there was no evidence for anti-sense DPRs. Anti-sense RAN translation either does not occur in these flies or levels are below the threshold for detection using these antibodies via dot blotting.

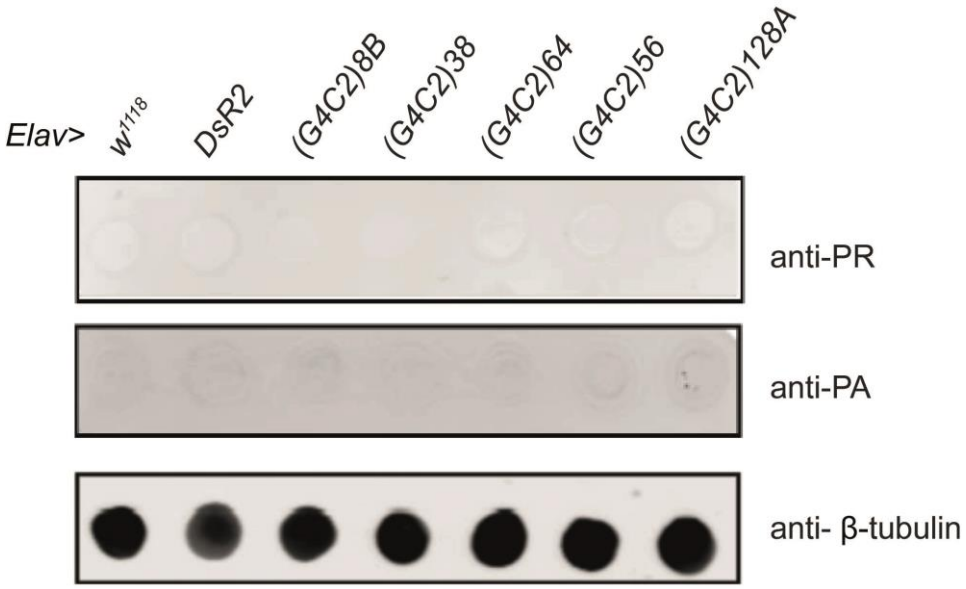

Figure 3.10 | Anti-sense DPRs were not detected in flies expressing G_4C_2 repeats. Published poly-PR and poly-PA antibodies (Gendron et al., 2013) were used to detect the presence of anti-sense DPRs in RIPA soluble lysate from flies expressing 38 G_4C_2 repeats. No anti-sense DPRs were detectable the using anti-PR and anti-PA antibodies. Samples were also probed with anti-β-tubulin to confirm approximately equal loading of protein.

3.10 Discussion

3.10.1 No RNA foci were detected in the L3 salivary glands of DsRed2-G_4C_2 flies

Using an Alexa488-labelled $(G_2C_4)_4$ RNA probe to perform FISH in the salivary gland cells of larvae expressing 32 or more repeats showed the presence of a single large "foci" per cell. Both Mizielinska et al. (2014) and Tran et al. (2015) previously observed a single large "foci" in salivary gland cells expressing G_4C_2 repeats. Tran et al. (2015) found these single large "foci" co-localised with activated Pol-II; indicating they represent the transcriptional start site. Indeed, when salivary gland cells expressing the repeat constructs used in this study were co-stained with the Alexa488-labelled $(G_2C_4)_4$ RNA probe and an antibody for p-Pol-II, the single foci observed co-localised with p-Poll-II. The lack of foci in the G_4C_2 repeat containing lines does not exclude the possibility that the foci are produced from the constructs used in this study. Indeed the 38 and 56 repeat constructs have been shown to form numerous nuclear RNA foci in SH-SY5Y cells previously by Lee et al. (2013). Tran et al. (2015) could detect numerous small and large foci in larval glia and neurons from their 160 intronic construct; hence it will be important in future to investigate RNA foci in neurons for the DsRed2-repeat flies used in this thesis. Tran et al. (2015) also studied G_4C_2 RNA transcribed as part of a poly(A) mRNA using 36 G_4C_2 repeat flies (36R-poly-(A)) generated in the Mizielinska et al. (2014) study. Not only were the levels of RNA in the 36R-poly-(A) flies much greater than in the spliced intronic 160 G_4C_2 flies despite the same transgenic insertion site, but most of the RNA in these flies was cytoplasmic in larval glia and motor neurons; with the only nuclear signal being the single large dot representing the transcriptional start site. The G_4C_2 repeats in the DsRed2 flies used in this study are in the 3'UTR so will be part of the poly-(A) mRNA transcript like the 36R-poly-(A) flies where most of the RNA is cytoplasmic (Mizielinska et al., 2014, Tran et al., 2015). Nuclear RNA foci might therefore potentially be difficult to detect in the current model due to the nature of the construct as most of the RNA is exported to the cytoplasm where it undergoes RAN translation. Any foci seen will likely be cytoplasmic like what is observed for the 36R-poly-(A) flies.

Importantly, in salivary gland cells expressing the 288 G4C2 RNA-only construct, although a large bight signal co-localising with p-Pol-II was seen, G_4C_2 RNA was also detected in the nucleus that didn't co-localise with p-Pol-II - including several smaller dots akin to the foci seem in patient iPSC neurons (Almedia et al., 2013). Hence this RNA only fly line was deemed an appropriate RNA-only control for later studies investigating DPR toxicity from alternative codon constructs in salivary gland cells.

3.10.2 DsRed2 levels are higher in DsRed2-G_4C_2 repeat containing flies

DsRed2 protein levels were found to be elevated in flies harbouring >8 G_4C_2 repeats in the 3'UTR of *DsRed2* compared to controls expressing only *DsRed2* without repeats in the 3'UTR. The 3'UTR of mRNA is important for the localisation and stability of the transcript (Matoulkova et al., 2012). The G_4C_2 repeat RNA forms DNA and RNA G-quadruplexes (G4) (Haeusler et al., 2014). RNA G4s in the both the 5' and 3' UTR of an RNA transcript have been demonstrated to have key functions in RNA metabolism including mRNA splicing, polyadenylation, mRNA localisation and importantly mRNA translation (Song et al., 2016). For example, using luciferase reporter assays the presence of 5' UTR G4 structures have been shown to correlate with translation repression of numerous human mRNAs (Song et al., 2016). Furthermore, Crenshaw et al. (2015) found a G4 in the 3′-UTR of *APP* mRNA inhibits the expression of the APP protein. Beaudoin and Perreault, (2013) found G4s in the 3' UTR can enhance translation by increasing the efficiency of alternative polyadenylation sites. Most evidence suggests 5' and 3' UTR G4s inhibit cap-dependent canonical translation via influencing eIF4A-sensitivity (Song et al., 2016). Intriguingly the opposite appears to be true for non-canonical cap-independent forms of translation, G4 structures in the 5'UTR enhance IRES mediated mRNA translation rather than supress it through unknown mechanisms (Song et al., 2016). Hence the enhanced DsRed2 levels seen in the presence of G_4C_2 repeats in the *DsRed2* 3'UTR could be due an interaction between RNA G4 structures formed by the repeats and proteins important non-canonical translation; if G_4C_2 RAN translation occurs through cap-independent processes.

3.10.3 Length dependent RAN translation

DPRs RAN translated from the G_4C_2 mRNA is one of the major unique pathological hallmarks of the C9ALSFTD brain. In this study DPRs were detectable in flies harbouring 32, 56 64 and 38 G_4C_2 repeats. No DPRs were detectable in flies harbouring shorter than 32 repeats. The threshold for RAN translation of the G_4C_2 expansion is yet to be determined however clues can be gathered from models of repeat toxicity and patients harbouring the repeat expansion.

Gami et al. (2015) found all 5 DPRs in the cerebellum, frontal and temporal cortices of a 30 unit G_4C_2 repeat expansion carrier. However, no DPRs are detected in post-mortem brain tissue from <20 G_4C_2 repeat unit cases. For example, Beer et al., (2014) identified a patient with clinically typical ALS with a 16 unit G_4C_2 repeat expansion with no DPR pathology. Hence patient post-mortem studies suggest the critical length for RAN translation lies somewhere between 20 – 30 repeats. Cellular models also suggest a critical repeat number is required for the initiation of RAN translation. Zu et al. (2013) detected poly-GP, poly-GA and poly-GR from a 30 G_4C_2 repeat construct transfected in HEK293T cells. RAN translation was again length dependent – no DPRs were observed in a 4 G_4C_2 repeat unit construct but DPR levels increased as the repeat length increased from 30 to 60 and 120 repeats. Mori et al. (2013a) also found a length-dependent increase in translation of the poly-GP and poly-GA peptides from G_4C_2 repeat constructs *in vivo*. Increasing amounts of poly-GA was detectable beginning at 38 repeats – no DPRs were detectable at 11 and 28 repeats. *Drosophila* models of the G_4C_2 expansion have also provided insights into the repeat length required for RAN translation to occur. DPRs have been observed at 28, 30, 36, and 58 repeat units but not at 3 and 8 repeats (Mizielinska et al., 2014, Freibaum et al., 2015, Zhang et al., 2015). Freibaum et al. (2015) observed minimal translation at 28 repeats with a large increase in detectable DPRs at 58 repeats. Zhang et al. (2015) expressing a 30 repeat unit construct found poly-GR was only detectable with 2x copies of the construct and required a heat-shock GAL4 with an accompanying heat-shock for expression. Poly-GR was not seen with a single copy of the 30 repeat construct, furthermore poly-GP was not observable under any conditions. However, Zhang et al. (2015) found both poly-GR and poly-GP were measurable in flies with a single copy of a 36 repeat construct

generated by Mizielinska et al. (2014); furthermore poly-GR did not require a heat-shock to be detected. Poly-GA aggregates are seen throughout the brain of BAC mice with 36 repeats; however, it appears expression levels of G_4C_2 RNA need to be sufficiently high enough as no DPRs were seen in mice with 37 repeats in the same study (Liu et al., 2016c). Similarly, the different fly models have differences in the transgenic insertion sight and genomic context of the repeats - both of which strongly influences transgene expression levels, hence it is difficult to directly compare between the different fly models.

Nevertheless, when fly (Mizielinska et al., 2014, Freibaum et al., 2015, Zhang et al., 2015) *in vitro* (Zu et al., 2013, Mori et al., 2013a) post-mortem human work (Beer et al., 2014, Gami et al., 2015) and BAC models (Liu et al., 2016c) are all viewed in parallel - in addition to the results presented in this study - a noticeable trend becomes apparent. It seems RAN translation of the G_4C_2 expansion initiates at a repeat unit number in the high 20's but is however likely weak and inefficient based on what has been observed (Mori et al., 2013a; Freibaum et al., 2015). However, at repeat lengths >30 repeats, there is large increase RAN translation efficiency and hence DPR production. In the current study a significant difference in poly-GP and poly-GA levels was seen for the 32 and 64 repeat lines, with the 64 repeat flies producing higher amounts of both DPRs. In all the studies discussed previously DPR expression increases as the G_4C_2 repeat number increases. For example, in the Liu et al. (2016a) BAC transgenic model mice with 500 repeats showed more severe DPR accumulation compared to those with 36 repeats. This finding that the initiation of RAN translation is strongly influenced by the length of the repeats is similar to findings on RAN translation in other neurological diseases caused by repeat expansions (Zu et al., 2011; Kearse et al., 2016). In summary studies of the hexanucleotide repeat expansion suggest the critical length for RAN translation of from the G_4C_2 expansion is in the range of 20 – 30 repeat units and that greater repeat length favours RAN translation and DPR accumulation.

It is also worth noting that both poly-GP and poly-GA ran at a higher molecular weight that would be predicted based on the repeat size from which they are translated from. Such a phenomenon is seen in other studies investigating poly-GP from pure and alternative codon repeat constructs (Zu et al 2013; Yamakawa et al., 2015; Mizielinska et al., 2014). Yamakawa et al. (2015) found poly-GA, poly-PA,

poly-PR and poly-GP all seem to run higher than expected indicating a decreased electromobility of highly repetitive peptides potentially due to abnormal SDS-binding properties, the way the DPRs fold, or a post-translational modification (Yamakawa et al., 2015). It is also possible the unique C-terminal sequences of poly-GP and poly-GA due to translation from the 3'UTR of DsRed2 may also influence how the proteins run on the gel. Multiple poly-GP bands were seen in the 32 and 64 repeat lines at different molecular weights; multiple poly-GA bands were seen in the 38 and 64 repeat flies. There are many reasons why this might occur - for example the proteins could be potentially cleaved, RAN translation could be initiating at different points along the repeat sequence or terminating at different points producing different sized poly-GP and poly-GA peptides. Indeed Tabet et al., (2018) and Green et al., (2017) both report codons 5' of the repeats other than the CUG in the (+1) frame also influences RAN translation albeit to a lesser extent than the (+1) CUG. If RAN translation is initating at different sites within the repeats then the higher than expcted molecular weight bands may be due to large C-terminal read-throughs of the DsRed2 3'UTR, hence it will be important to determine extact start site of RAN translation in the repeat constructs used in this study through either ribosomal profiling or site directed mutagenisis. Another explanation for multiple bands seen on SDS-PAGE for poly-GP and poly-GA is that the peptides may self-associate to form dimers or even higher order oligomers and not just exist in a monomeric state. Multiple band sizes for poly-GP by western blotting were reported by Yamakawa et al., (2015) and Cristofani et al., (2017) using a 100 peptide repeat alternative codon poly-GP only construct.

3.10.4 Construct dependent RAN translation

Both the expression and identity of the DPRs detected in the different G_4C_2 fly lines used in this study varied dramatically in a length and construct dependent manner (Table 1). The 32 and 64 repeat lines produce high levels of poly-GP and poly-GA whilst the 56 repeat line also produces only poly-GP and poly-GA but at much lower levels. The 38 repeat line produces all three sense DPRs, poly-GP, poly-GA and poly-GR. All constructs were inserted at the same location in the genome in the 3'UTR of the DsRed2 transgene; hence differences in DPR expression profile cannot be accounted for by genomic insert.

Table 1 | The different linkers 5' of the repeats in relation to DPRs detected and DPR expression levels. Constructs with the same linker sequence (32, 64 repeat constructs) produce the same combination and levels of DPRs.

# G4C2 repeats	Expression Level	DPR Type	DPR Levels
≤ 8	n/a	none	n/a
56	+	GA/GP	+/++
32	+	GA/GP	+++/+++
64	+	GA/GP	+++/+++
38	++	GA/GP/GR	++/+/+

*Expression level based on quantitative western blotting of DsRed2 protein levels; DPR levels based on quantification of DPR levels by western blotting.

The G_4C_2 repeats in the flies used in this study are contained within the DsRed2 poly(A) mRNA transcript, hence are likely exposed to ribosomal scanning during DsRed2 translation. No translational read-through of DsRed2 is seen (figure 3.3); however, it is known that the ribosomal complex can keep scanning after translation of the upstream open reading frame and reinitiate again further downstream (Jungreis et al., 2011; Rogers et al., 2017). Hence, DPR translation is initiating either in the sequence 5' of the repeats and/or within the repeat sequence itself. The repeats are uninterrupted in these flies and identical within each construct apart from length, however there is variability in sequence 5' of the repeat region. Indeed, when generating these constructs, some variation in the linker sequence between the DsRed2 stop codon and the start of the G_4C_2 repeat sequence was introduced by chance. The 5' linker flanking the start of the repeats, after the DsRed2 STOP codon, differs in between the 38, 56, 32 and 64 repeat constructs (figure 3.11).

Recent work has begun shed light upon potential mechanisms underlying RAN translation. Kearse et al. (2016) demonstrated RAN translation of CGG repeats in fragile X-associated tremor/ataxia syndrome requires traditional cap-dependent 40S ribosomal scanning, however the start site for translation differs from traditional AUG initiated translation. RAN translation, depending on the reading frame, could initiate within the repeat sequence itself or upstream of the repeats on AUG like codons. The addition of STOP codons upstream of the CGG repeat strongly inhibited RAN translation of the FMR-polyGly protein. Further, using mass spectrometry Sellier et

al. (2017) found translation of the FMR-polyGly protein initiated at an ACG close cognate start codon 32 nucleotides upstream of the CGG expansion. Investigations into mechanisms of RAN translation of G4C2 repeat expansion have revealed similiary that the translation of the DPRs is dependent on the sequence 5' of the repeats. G4C2 RAN translation initiates in the poly-GA frame (+1) on a near-cognate CUG codon 5' of the repeats in the context of a strong kozak sequence (Green et al., 2017; Tabet et al., 2018). Translation in poly-GP (+2) and poly-GR (+3) frames occurs through ribosomal frameshifting (Tabet et al., 2018).

In the DsRed2-G4C2 model characterized in this study DPR levels and production correlated with the sequence 5' of the G4C2 repeats. For constructs with the same linker sequence and linker length the same pattern of DPR expression is seen; e.g. for the 32x and 64x repeat lines high levels of poly-GP and poly-GA are produced. The 56x repeat line on the other hand, despite a similar repeat length number to the 64 line, produces significantly lower levels of poly-GP and poly-GA - the linker sequence however is not only different but also longer in length (34bp compared to 24bp). Hence it is plausible to speculate that the different upstream sequences 5' of the repeats in the different DsRed2-repeat flies' impacts RAN translation. However, analysis of the codons within these linker sequences reveals DPR levels and identity does not correlate with the presence of near-cognate start codons or STOP codons (figure 3.11). A CUG is present in the (+1) frame 5' of the repeats in all DsRed2 constructs however, the presence of this CUG does not explain the DPR expression profile observed in these flies. For example. In the 32 and 64 repeat flies the CUG codon is present in the (+1) poly-GR reading frame, yet poly-GR is not detected in these lines. Rather poly-GA (+2) and poly-GP (+3) are produced much more abundantly. The 38 repeat flies, the only line in which poly-GR is detected harbour two AUG like (AUC and AGG) codons in the (+2) poly-GR frame, however a UAG STOP codon is present in the poly-GA (+3) frame just before the repeats yet this DPR is detectable in these flies.

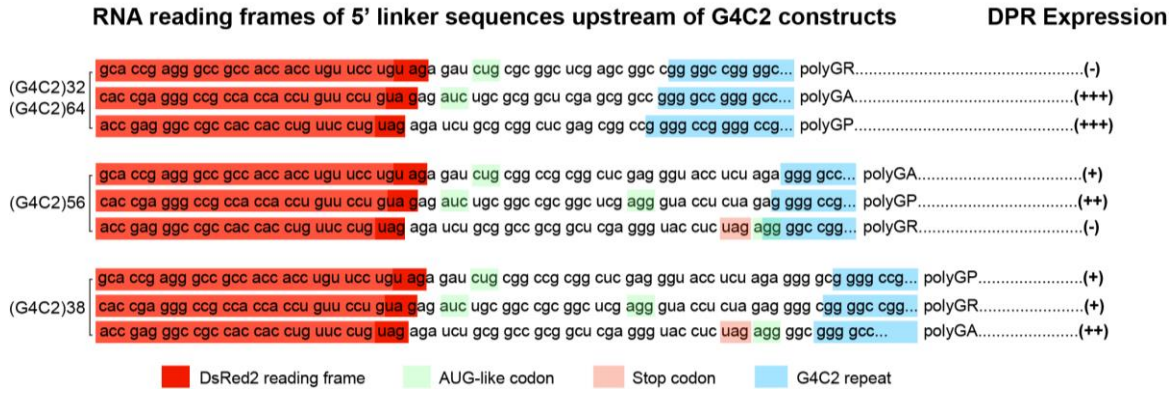

Figure 3.11 | AUG-like and STOP codons in the 5' linker sequences of the G4C2 DsRed2 flies that produce DPRs. Schematic of constructs with different lengths of uninterrupted G4C2 repeats cloned into 3' UTR of disease-unrelated marker gene DsRed2 (in red); variable linker sequences between DsRed2 stop codon and start of G4C2 repeats are indicated, as are AUG-like codons and DPR produced. Expression levels (+, low; ++, moderate; +++ high) of RNA-translated DPR produced by specific G4C2 repeat construct.

The different DPR profiles in the flies used in this study are an interesting observation; the current study however is limited in its ability to draw conclusions about mechanisms of RAN translation other than speculation. To study whether the 5' linker sequence is behind the different DPR profiles observed; further work should be conducted where the linkers are swapped between different constructs and DPR expression profile then probed for. For example, if the 56x G_4C_2 construct has the same 5' sequence upstream of the repeats as the 32 and 64 lines would it now produce high levels of poly-GP and poly-GA? Similarly, if the 32, 56 or 64 repeat constructs have the same 5' linker as the 38, would they start to produce poly-GR?

3.10.5 No anti-sense DPRs are detectable in DsRed2-G_4C_2 repeat lines

The anti-sense DPRs poly-PR and poly-PA were not detected in any of the flies used in this study using previously published antibodies. So far anti-sense foci, but not anti-sense DPRs, have been reported in any of the published G_4C_2 mouse models (Chew et al., 2015; Peters et al., 2015; O'Rourke et al., 2015; Liu et al., 2016c; Jiang et al., 2016). Anti-sense DPRs have also not been reported in other fly models of the repeat expansion (Mizielinska et al., 2014; Freibaum et al., 2015; Zhang et al., 2015;

Tran et al., 2015). How anti-sense transcription of the C_4G_2 sequence occurs is unknown. Anti-sense transcripts are known to arise from independent promoters, bidirectional promoters and cryptic promoters (Pelechano and Steinmetz, 2013). Hence it is likely some form anti-sense promotor is present within the *C9ORF72* gene. If this is the case then no anti-sense DPRs or even anti-sense RNA would be seen from model systems that study the repeat expansion in isolation of the full *C9ORF72* genomic landscape such as viral vector mouse models (Chew et al., 2015) or *Drosophila* models (Mizielinska et al., 2014; Freibaum et al., 2015; Zhang et al., 2015; Tran et al., 2015).

There are several explanations as to why no anti-sense DPRs were observed. It could be that their expression levels are too low to be detectable via dot blotting; indeed poly-PR and poly-PA are extremely rare in the human brain (Mackenzie et al., 2015); indicating anti-sense RAN translation is less efficient than sense RAN translation. It is also possible anti-sense DPRs are not produced at all as anti-sense transcription requires some form of promotor (Pelechano and Steinmetz, 2013), hence anti-sense RNA would not be produced from the UAS/GAL4 system used to express the repeats in *Drosophila*. Furthermore, the presence of bidirectional promoters appears to be species-dependent, with the genome of *Drosophila* containing much lower levels of bidirectional promoters compared to yeast and humans (Pelechano and Steinmetz, 2013). To produce anti-sense RNA the deliberate insertion of an anti-sense UAS for the GAL4 transcription factor to bind to would be needed. Further one could also utilise a C_4G_2 construct under the control of a UAS promotor. Transgenic BAC models with the whole 29.6kb *C9ORF72* gene, with upstream and downstream sequences, potentially provide the means to study both sense and anti-sense DPRs in the same model system as the whole *C9ORF72* genomic landscape is present. Anti-sense foci have been detected in *C9ORF72* BAC mice containing 140.5kb upstream and exons 1-6 of *C9ORF72* (Peters et al., 2015) and *C9ORF72* BAC mice containing 140kb upstream and exons 1-5 of *C9ORF72* (Jiang et al., 2016) further suggesting some form of anti-sense promotor lies within this region. Although, no anti-sense RNA was detected in the intronic Tran et al. (2015) *Drosophila* model consisting of exon 1, some of intron 1 (with the G_4C_2 expansion) and exon 3 of human *C9ORF72*. Human *C9ORF72* iPSCs also give the same benefits as BAC transgenic mice as the whole *C9ORF72* gene with

the repeats can be studied, however, anti-sense foci or DPRs have not been fully investigated in current *C9ORF72* iPSC lines (Donnelly et al., 2013; Sareen et al., 2013; Lopez-Gonzalez et al., 2016; Dafinca et al., 2016); although trace amounts of poly-PR were reported by Dafinca et al. (2016) in their C9 iPSC motor neurons.

3.10.6 Poly-GP and poly-GA localisation differed in different DsRed2-G_4C_2 repeat lines

The localisation of poly-GP varied substantially between the 56 and 64 repeat lines. In the 56 repeat line poly-GP was mainly diffusely nuclear with intense nuclear poly-GP accumulations consistent with the paranucleolar inclusions seen in C9ALS/FTD frontal cortex (Schludi et al 2015). For the 64 repeat flies poly-GP was predominantly diffusely cytoplasmic with a weak nuclear stain but with no intense nuclear accumulations.

The cytoplasmic inclusions of poly-GP seen in patient post-mortem brain were not observed in any of the flies used in this study. Numerous cellular models used to assess poly-GP toxicity using both G_4C_2 and alternative codon constructs have found the peptide to be predominantly diffusely localised in both the cytoplasm and nucleus without the formation of inclusions (Zu et al., 2013; May et al., 2014; Zhang et al., 2014; Yamakawa et al., 2015; Schludi et al., 2015; Tao et al., 2016; Callister et al., 2016). Nuclear and cytoplasmic inclusions of poly-GP have been reported in one *Drosophila* model expressing 58 G_4C_2 repeats (Freibaum et al., 2015) and in aged G_4C_2 repeat murine models (Chew et al., 2015, Peters et al., 2016, O'Rourke et al., 2016, Liu et al., 2016c, Jiang et al., 2016). However numerous cellular models used to assess poly-GP toxicity using both G_4C_2 and alternative codon constructs have found the peptide to predominantly diffusely localised in both the cytoplasm and nucleus without the formation of inclusions (Zu et al., 2013; May et al., 2014; Zhang et al., 2014; Yamakawa et al., 2015; Schludi et al., 2015; Tao et al., 2016; Callister et al., 2016). This is consistent with the biophysical properties of the peptide that it is uncharged with a compact flexible coil structure (Freibaum and Taylor., 2017). Poly-GP is much for soluble than the most aggregation prone and common DPR poly-GA, indeed a greater percentage of total poly-GP is soluble compared to insoluble in brain tissue whilst the opposite is the case for poly-GA (Gendron et al., 2015). This solubility of poly-GP has led to investigations in the potential of the peptide as

biomarker for C9ALS/FTD in CSF (Gendron et al., 2017; Lehmer et al., 2017). Hence poly-GP does not appear to be highly prone to forming insoluble proteinaceous inclusions on its own. Poly-GP inclusions may occur due to its interaction with other DPRs. For example, Yamakawa et al. (2015) found poly-GR and poly-PR sequesters both poly-GP and poly-GA into aggregates. Lee et al., (2017) found poly-GA sequesters poly-GP. Poly-GR has been shown to accumulate in cytosolic poly-GA inclusions in Drosophila and human iPSC neurons (Yang et al., 2015). Poly-GP aggregates could be age-dependent and attributable to a build-up over time of the peptide in the patient brain. A recent study demonstrated that whilst under basal conditions poly-GP mainly accumulates as an SDS-soluble species levels insoluble poly-GP increase 5 fold when autophagosome formation is inhibited (Cristofani et al., 2017). The authors also found poly-GP was the only DPR to accumulate in the insoluble fraction following proteasome inhibition. The C9ORF72 protein itself has been shown to play a role in the initiation of autophagy (Webster et al., 2016a; Yang et al., 2016) and its loss impairs autophagic functions (Sellier et al., 2016). Further poly-GA toxicity has been linked to proteasome dysfunction (May et al., 2014; Zhang et al., 2014; Zhang et al., 2016). The interaction between both factors may be important for the aggregation of poly-GP.

Many model systems studying poly-GA toxicity in isolation have shown the peptide to be highly aggregation prone – predominantly forming cytoplasmic inclusions (Zu et al., 2013; May et al., 2014; Zhang et al., 2014; Yamakawa et al., 2015; Schludi et al., 2015; Tao et al., 2016; Callister et al., 2016). Poly-GA was undetectable in the 56 repeat line but interestingly was predominantly nuclear in the 64 repeat line; forming nuclear inclusions. As was the case for poly-GP in the 56 repeat flies, these nuclear poly-GA inclusions seen in the 64 repeat line strongly resembles the paranucleolar poly-GA inclusions seen in C9ALS/FTD frontal cortex (Schludi et al., 2015). Nuclear inclusions of poly-GA have been observed in a poly-GA protein only mouse model where it sequestered proteins involved in nuclear cytoplasmic transport and nuclear UPS-related proteins (Zhang et al., 2016). The nuclear localisation of poly-GP and poly-GA may be attributable to the small size of the dipeptides; in mice poly-GP produced 66 repeats is primarily nuclear (Chew et al., 2015) whereas poly-GP forms mainly cytoplasmic aggregates in mice with repeat expansions in the hundreds (O'Rourke et al., 2015; Peters et al., 2015; Jiang et al 2016). Jiang et al. (2016)

suggest shorter DPR proteins are more soluble hence more freely diffuse through the nuclear pore forming the nuclear aggregates seen by Chew et al. (2015).

The difference in poly-GP localisation between the 56 and 64 flies could be attributable to the vastly different expression levels of poly-GP between these two lines. A second explanation may be due to the different non-GP peptide C-terminals the poly-GP proteins are predicted to have in the 56 and 64 repeat lines assuming translation halts at the first stop codon 3' of the repeats (figure 3.12). These C-terminals may alter the biophysical properties of the peptide. Interestingly all six RAN proteins are predicted to have unique C-terminal sequences in the human brain (Zu et al., 2013). Immunostaining with specific antibodies has shown DPRs with unique C-terminal flanking regions aggregate in patient brains (Zu et al., 2013). How these C-terminals influence DPR biophysical properties or toxicity is unknown but broadens the study of DPR toxicity further than just the repetitive peptide motifs themselves (Cleary and Ranum, 2017). Although it should be noted that the C-terminal sequence for the 56 and 64 repeat flies do not differ greatly, for example, the predicted C-terminal for poly-GP in the 56 repeat flies is identical to the C-terminal for the 64 repeat flies except for an extra Proline. To confirm/exclude whether the different subcellular localisation of poly-GP and poly-GA between the repeat flies is due to the extra peptide sequence translated from the 3'UTR of DsRed2, alternative codon constructs could be generated for poly-GA and poly-GP which harbour these extra C-terminal amino acids. The expression of such constructs could then be compared to poly-GA and poly-GP constructs without this C-terminal to see if localisation differs.

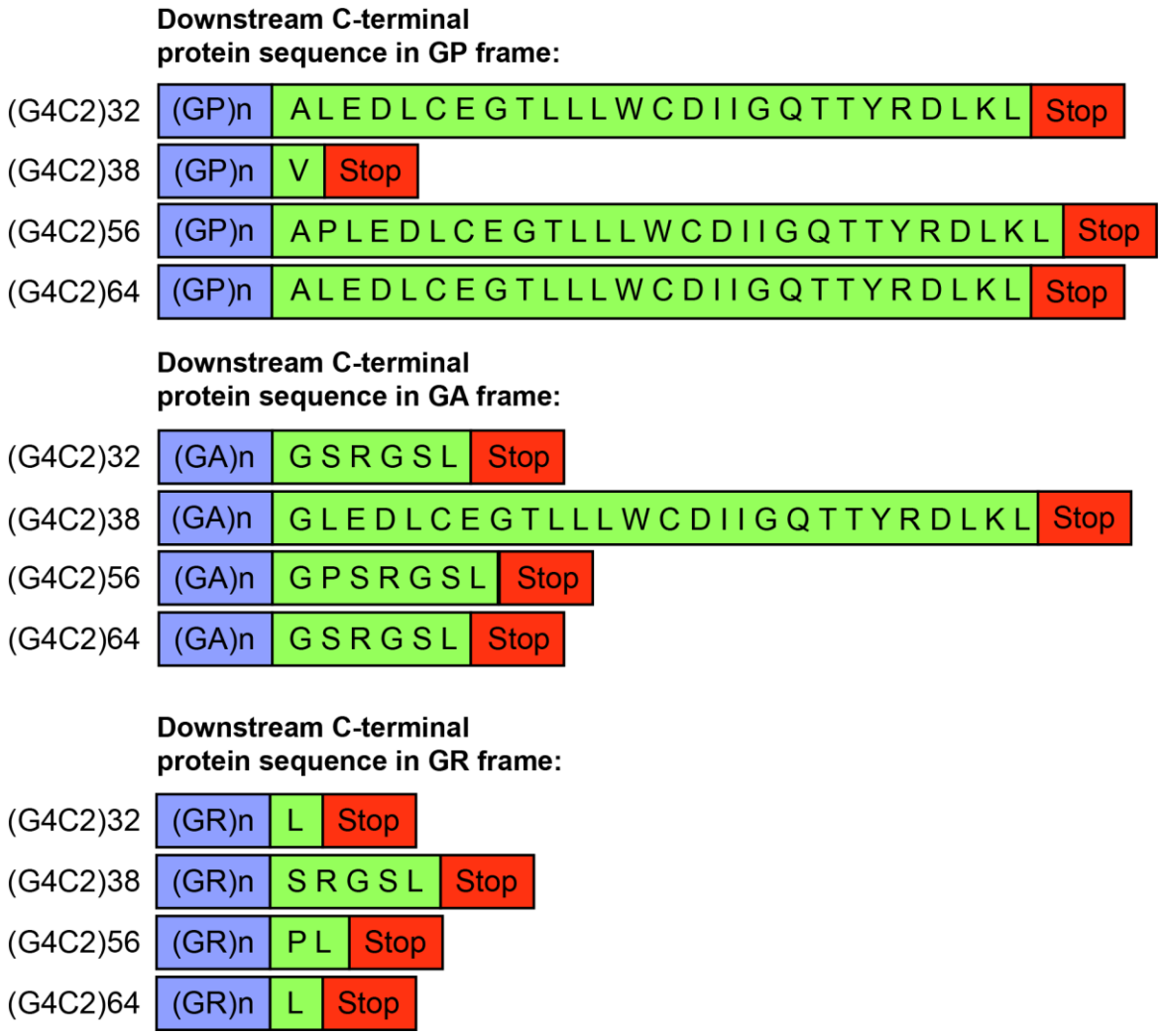

Figure 3.12 | Construct specific differences in the putative C-terminal sequences of DPRs. The C-terminal protein sequences between the repeats and the first downstream stop codon are predicted to differ in length and amino acid composition for the different constructs in each reading frame. This figure was generated by Alan Stepto and used with his permission.

3.10.7 Conclusion

In summary this chapter has demonstrated RAN translation occurs in a novel *Drosophila* model of the repeat expansion in which the repeats have been cloned into the 3'UTR of the disease unrelated transgene *DsRed2*. Only RAN translation of the sense DPRs were observed; no anti-sense DPRs were detected. No DPRs were seen in flies with 8 repeats however were seen from 32 repeats and above. DPR identity and expression differed between the different G_4C_2 fly lines. The 32 and 64 repeat fly lines produced poly-GP and poly-GA. The 56 repeat fly line also produced only poly-GP and poly-GA but at a lower level compared to 32 and 64 flies. Poly-GR is detectable only in the 38 repeat fly line. Hence RAN translation occurs in this model, however in both a length and construct dependent manner; potentially due to construct specific differences in the linker sequence 5' of the repeats or construct specific differences in the putative C-terminal sequences of the DPRs.

Chapter 4: Neuronal toxicity of G_4C_2 repeats is dependent upon both DPR identity and expression levels

The preceding chapter described DPR expression and localisation in flies harbouring different lengths of G_4C_2 RNA cloned in the 3' UTR of *DsRed2*. RAN translation occurs in this model, however in both a length and construct dependent manner. The 64 repeat line produces high levels of poly-GP and poly-GA compared to the 56 repeat line - despite similar G_4C_2 repeat length sizes. RAN translation of poly-GP and poly-GA is also observed in the 38 repeat flies, however uniquely also produces the third sense DPR, poly-GR. These reported differences in DPR levels and expression profile between different G_4C_2 expressing fly lines facilitated assessment of the toxicity of different combinations and levels of DPRs derived from RAN translation of G_4C_2 repeat RNA in the following chapter.

4.1 Neuronal expression of 38 repeats produces a severe early-onset deficit in climbing performance whereas late-onset decreases in climbing performance are seen in flies expressing high levels of poly-GP and poly-GA

C9ALS/FTD patients are characterised by a progressive decline in motor functions resulting from the loss of lower and upper motor neurons (Swinnen and Robberecht, 2014). Whether the expression of G_4C_2 RNA and RAN translated DPRs had any impact on *Drosophila* motor performance was investigated. Assessment of motor behaviour resulting from neuronal expression of G_4C_2 repeats was performed by measuring the innate climbing response of flies through startled induced negative geotaxis (SING). The natural response of a fly when startled in a closed environment is to climb up to the top of its enclosure. Defects in climbing performance are indicative of underlying neuronal dysfunction (Coulom and Birman, 2004). Climbing performance was investigated at days 5 and days 20 post eclosion using the pan-neuronal driver $Elav^{C155}$ (figure 4.1). At 5 days of age, flies expressing 38 G_4C_2 repeats recorded a significant and severe deficit in climbing ability, relative to all other genotypes tested (figure 4.1A and C). The climbing performance of the 38 repeat flies and its 8 repeat control at day 5 (*$Elav^{C155}$>$(G_4C_2)8B$*) were assessed in the same geotaxis assay as the two controls (*$Elav^{C155}$*/+ and *$Elav^{C155}$>DsR2*) and the other repeat lines; however, the data is separated for clarity in figure 4A and C. By day 20, the 38 repeat flies were unable to climb, whereas all other genotypes performed as they had at day 5 (figure 4.1B and D). As with the day 5 geotaxis data the climbing performance of all repeat flies and controls was assessed in the same geotaxis assay for day 20, however, the data set is again separated in the figure for clarity. In summary the 38 repeat line which produces all three sense DPRs, poly-GP, poly-GA and poly-GR have a severe early onset behavioural phenotype. This phenotype was not caused by the enhanced G_4C_2 RNA levels or increased transgene expression of DsRed2 in the 38 repeat line. This was confirmed by generating flies harbouring two copies of the 64 repeat construct which have significantly higher levels of DsRed2 protein compared to both flies with a single copy of the 64 construct and the 38 repeat flies (appendix 2, figure A2 A; performed

by Alan Stepto). Day 5 climbing performance was not significantly different between the 2x 64 flies and the 1x 64 flies; however, both were significantly impaired compared to the 38 repeat line (appendix 2, figure A2 B; performed by Alan Stepto). No difference in climbing performance was seen between the 32, 56 and 64 lines producing only poly-GP and poly-GA at days 5 and 20 between each other and controls.

To establish whether expression of poly-GA and poly-GP contribute to behavioural impairment with age in flies, the 32, 56 and 64 repeat flies and the respective controls were aged to 40 days and subjected to SING. The 38 repeat flies did not survive up to day 40 hence could not be assessed at this later time point. In SING the performance of flies expressing 56 G_4C_2 repeats with low levels poly-GA and poly-GP could not be distinguished from control flies expressing either 0 (*Elav/+* and *Elav^{C155}>DsR2*) or 8 G_4C_2 (*Elav^{C155}>(G_4C_2)8A*) repeats up to day 40 (figure 4.1E). Hence neither the 56 G_4C_2 repeat RNA nor the poly-GP and poly-GA expression in this line induced any measurable deficit in climbing performance. The 32 and 64 repeat lines however developed a mild, yet significant impairment in climbing performance compared to the 56 line and controls at day 40 (figure 4.1E). DsRed2 transgene expression levels did not differ in these between the 32, 56 and 64 lines (appendix 1: figure A1), indicating the mild motor impairment seen is attributable to the high poly-GP and/or poly-GA protein levels produced. Geotaxis was performed in collaboration with Alan Stepto.

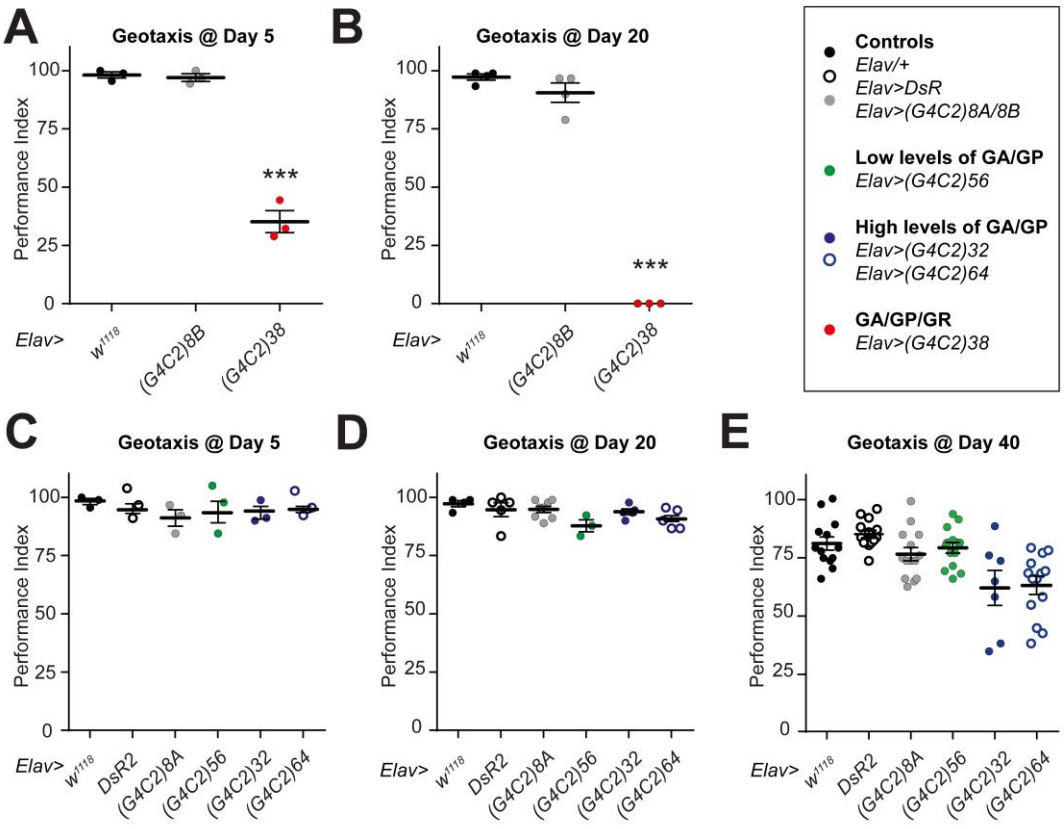

Figure 4.1 | A severe decline in climbing performance is seen in flies expressing 38 repeats expressing detectable levels of poly-GR; whereas late-onset decreases in climbing performance are seen in flies expressing high levels of poly-GP and poly-GA. **(A-D)** A two-way ANOVA was used to compare the mean differences in climbing performance (the performance index - PI) between the different genotypes at days 5 and 20. The PI of flies was significantly affected by both genotype ($F=183.5$, $p<0.0001$) and age ($F=21.61$, $p<0.0001$), further there was a significant interaction between genotype and age ($F=9.710$, $p<0.0001$). Post-hoc analysis was performed using Bonferroni-Holm corrected multiple comparisons which can be found in appendix 10, table A10. Post-hoc analysis revealed the 38 repeat flies had a significantly worse PI compared to all genotypes at day 5 ($p<0.0001$) and day 20 ($p<0.0001$). There were no significant differences for any other genotypes at these two time points. The performance of the 38 repeat flies and its respective controls at days 5 and 20 geotaxis was assessed in the same assay as the other genotypes but is separated in the figure for clarity. Each data point represents the PI of ~15 flies. Mean with SEM is shown for each genotype (n=3-4). **(E)** A significant difference in the PI of genotypes tested at day 40 was observed using a one-way ANOVA ($F=8.479$, $p<0.0001$). Post-hoc analysis was performed using Bonferroni-Holm corrected multiple comparisons which can be found in appendix 10, table A10. Post-hoc analysis revealed the 32 and 64 repeat flies had a significantly worse PI at day 40 compared to other genotypes. Note that flies expressing 32 and 64 repeats, but not 56 G4C2 repeats, show motor impairment by day 40 compared to controls**.** Each data point represents the PI of ~15 flies. Mean with SEM is shown for each genotype (n=7-14).

4.2 Open-field motor activity is severely impaired in the 38 repeat flies

As SING is a stimulus-dependent assay, flies were subjected to a stimulus-independent open-field motor assay to establish whether behavioural impairment persisted in this condition. Open-field motor tracking for the 38 repeat flies and the respective controls (*Elav/+* and *Elav(G4C2)8B*) were performed by Alan Stepto. Day 20 aged flies expressing 38 G_4C_2 repeats performed significantly worse than age matched controls expressing 0 or 8B repeats (figure 4.2A and B). The 38 repeat flies exhibited significantly reduced activity over a 30 minute recording period.

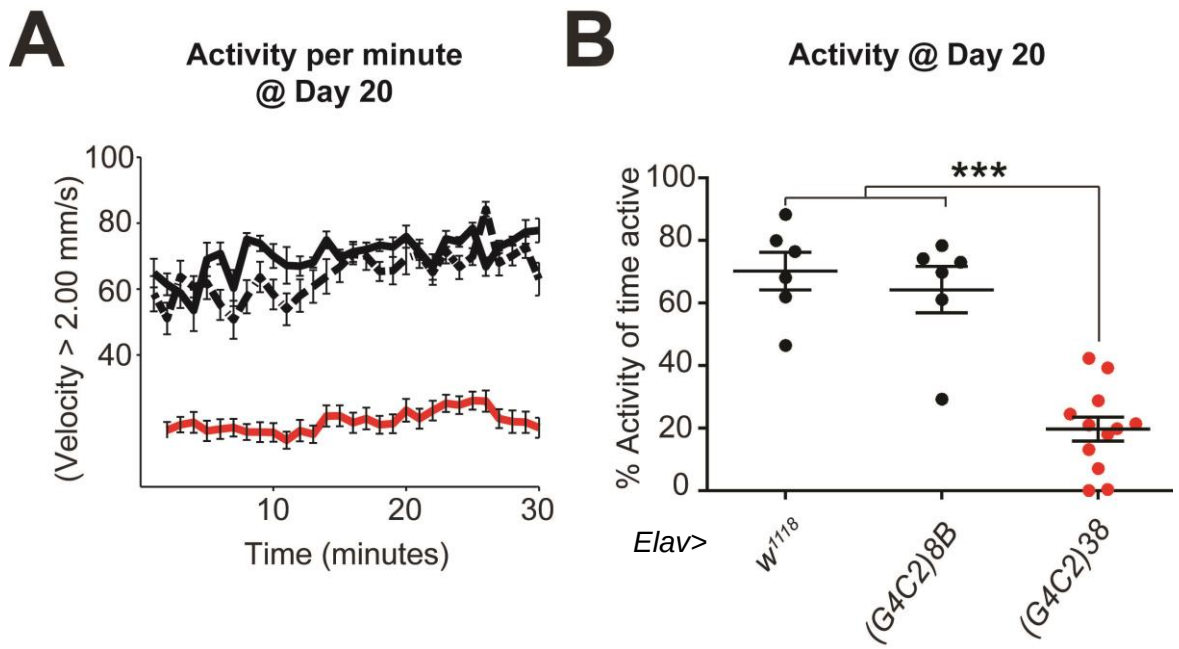

Figure 4.2 | Reduced open field activity is seen in flies expressing 38 repeats. (A) Open field behavioural activity at day 20; flies harbouring 38 repeats were less active than controls across the 30 minutes of tracking. **(B)** The percentage of the time each fly spent active was plotted. A one-way ANOVA revealed a significant effect of genotype on % activity of flies ($F=30.54$, $p<0.0001$). Post-hoc analysis was performed using Bonferroni-Holm corrected multiple comparisons which can be found in appendix 10, table A10. Post-hoc analysis revealed the 38 repeat flies had a significantly less % of time active compared to the controls (*Elav/+* and *Elav>(G4C2)8B*, $p<0.0001$) which did not significantly differ from one another. The data points represent the mean % activity for individual flies. Overall mean with SEM is shown for each genotype (n=7-14).

4.3 Open-field motor activity is impaired in flies expressing high levels of poly-GP and poly-GA at day 40

Stimulus-independent open-field motor performance was also conducted on day 40 flies expressing 8, 32 64 and 64 repeats in neurons. Analysis of the movement of these flies showed that the 32 and 64 G_4C_2 repeat lines, both of which produce high levels of poly-GP and poly-GA, have significantly decreased time spent active as compared to the DsRed2 (*Elav ^{C155}>DsR2*) and 8 repeat (*Elav^{C155}>(G_4C_2)8A*) controls and the 56 repeat line – which produces low levels of poly-GP and poly-GA compared with the 32 and 64 flies (figure 4.3A and B). Rasta plots for each genotype in which each individual bar represents the movement of a single fly over 30 minutes (with black space indicating when the fly is active and white space representing inactivity) were generated (figure 4.3C). These plots show that at the individual fly level the 32 and 64 flies spend much more time inactive than the 56 line and the 8 repeat and DsRed2 controls. Taken together with the SING data, when the high poly-GP and poly-GA 32 and 64 repeat lines are compared with the low poly-GP and poly-GA 56 repeat line - these results indicate that poly-GP and/or poly-GA when expressed at high enough levels cause an impairment in motor performance and by proxy neuronal dysfunction. Furthermore, as no motor impairment in the 56 repeat line is seen despite similar transgene expression levels (as measured by DsRed2 protein levels) between the 32, 56 and 64 lines - this suggest the expression of G_4C_2 RNA does not cause this toxicity consistent with other fly models studying G_4C_2 RNA toxicity *in vivo* (Mizielinska et al., 2014; Tran et al., 2015).

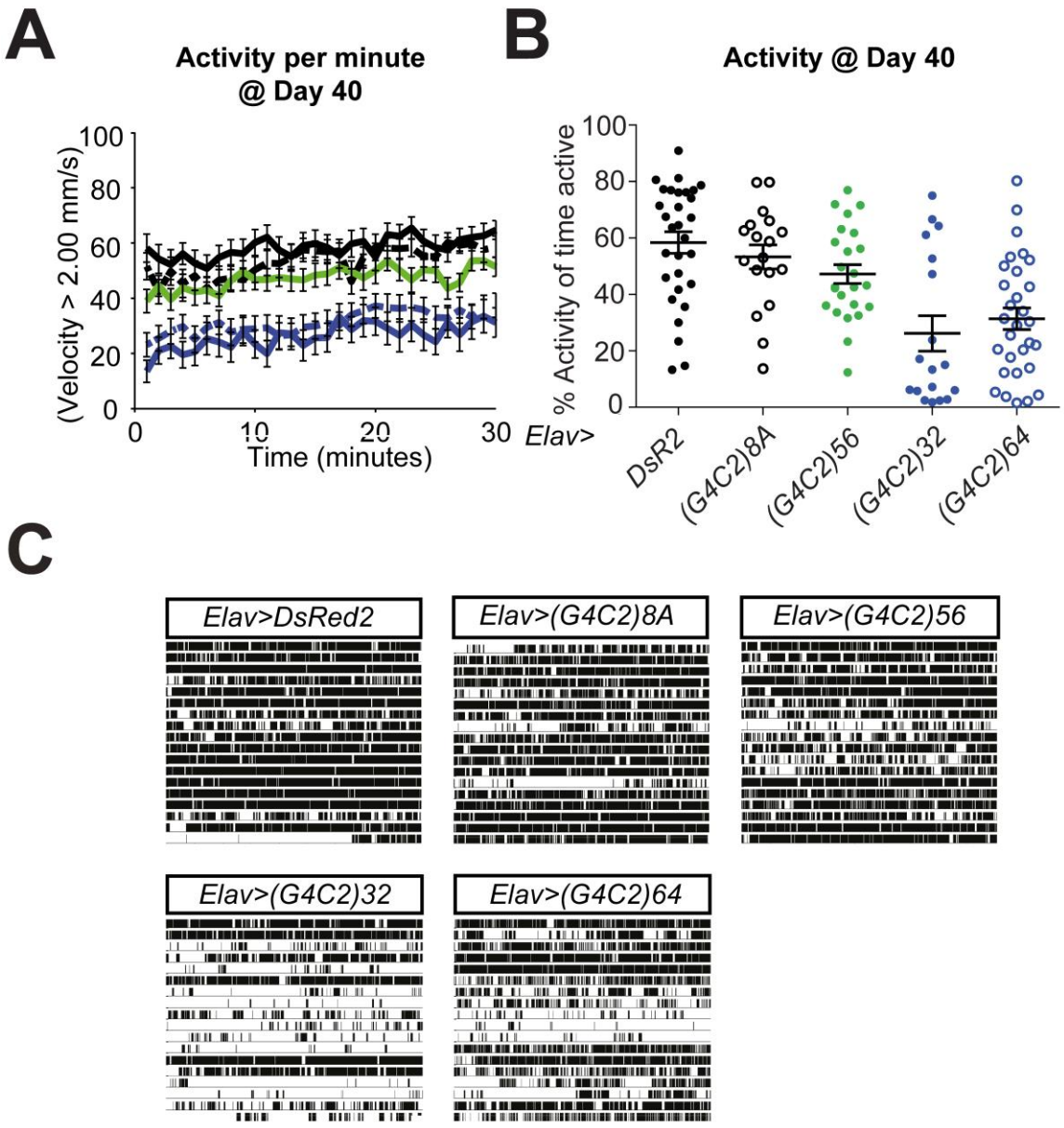

Figure 4.3 | Reduced open field activity is seen in aged flies expressing high levels of poly-GP and poly-GA. **(A)** Open field behavioural activity at day 20; flies harbouring 32 and 64 repeats were less active than controls across the 30 minutes of tracking. **(B)** The percentage of the time each fly spent active was plotted. A one-way ANOVA revealed a significant effect of genotype on % activity of flies ($F=10.77$, $p<0.0001$). Post-hoc analysis was performed using Bonferroni-Holm corrected multiple comparisons which can be found in appendix 10, table A10. Post-hoc analysis revealed flies expressing 32 and 64 repeats were significantly less active than control and flies expressing 8 or 56 repeats. Data points are mean % activity for individual flies with overall mean and SEM (n=18-30). **(C)** Raster plots demonstrating activity of individual flies over a 30 minute period. White bars represent periods of inactivity whereas black bars represent bouts of activity. Flies with 32 or 64 G_4C_2 repeats that express high levels of poly-GP and poly-GA were less active compared to controls and flies with 56 G4C2 repeats that express much lower levels of poly-GP and poly-GA.

4.4 Age dependent neurodegeneration is enhanced in flies expressing 38 G_4C_2 repeats

The most severe behavioural phenotypes correlate with the presence of poly-GR, whereas poly-GA and poly-GP correlate with mild later onset deficits in motor performance. This data indicates poly-GR is a highly potent driver of toxicity in *Drosophila* neurons whereas poly-GP and poly-GA are relatively well tolerated until later ages. To explore whether the severe motor impairment seen in the poly-GR expressing 38 repeat line was caused by underlying neuronal abnormalities age-dependent neurodegenerative cell loss was investigated. The compound *Drosophila* external eye and underlying photoreceptor morphology is classical way to assess neurodegenerative phenotypes in flies (Hirth, 2010). The DsRed2 control, 8, 64 and 38 repeat constructs were expressed in the largest class of photoreceptor cells in the *Drosophila* compound eye using *Rh1-GAL4,* the promotor for the major rhodopsin in the fly eye (Colley, 2012). The adult *Drosophila* compound eye consists of around 800 ommatidia arranged hexagonally (Kumar, 2012). Each hexagonal ommatidium contains 8 photoreceptors (R1-R8), the R1-R6 photoreceptors sit on the outside of the hexagonal ommatidium where as R7 and R8 are in the centre - the R7 photoreceptor sits atop R8, hence in any given focal point of view only R1-R7 photoreceptors can be viewed (Kumar., 2012). The R1-R7 photoreceptors in an ommatidium can be easily seen under a bright field microscope and the number of photoreceptors in an individual ommatidia counted and scored (e.g. % of ommatidium with 7, 6, 5 etc. photoreceptors). In Drosophila photoreceptors are neurons which send axonal projections to structures within the fly brain (Mollereau et al., 2001) hence the the compound eye is commonly used to easily investigate neurodegenerative phenotypes (Hirth, 2010)

Semithin retinal sections and quantification were performed in collaboration with Adel Boudi (Fanto Lab, KCL). As Drosophila photoreceptors are neurons, neuronal loss was measured by the number of remaining photoreceptor neurons in each ommatidium at 35 days of age. No observable differences in R1-R7 photoreceptor were seen between the controls and the G_4C_2 repeat containing lines after 14 days (data not shown). At 35 days of age some degeneration was seen in DsRed2 control and the 8B and 64 repeat lines (figure 4.5A). The representative distribution of

neurodegeneration in the fly retina was plotted as histogram for each genotype with number of photoreceptors per ommatidium (1-7) plotted on the X axis and % of total ommatidia with a specific number of photoreceptors on the Y axis (figure 4.5B). For the DsRed2, 8B and 64 fly lines the majority of ommatidium had between 4-6 photoreceptors (figure 4.5B). However, for the 38 repeat line most of the ommatidium contained between 1-3 photoreceptors. Hence enhanced age-dependent neurodegeneration of photoreceptors was seen only in the 38 repeat line when compared to the controls (*Rh1>DsRed2* and *Rh1>(G_4C_2)8B*) and 64 G_4C_2 repeat fly line. 100 ommatidium were counted per individual compound eye (N of 5-6 eyes for each genotype). To calculate the total % degeneration seen in the 38 repeat line compared to the DsRed2, 8B and 64 lines the same 100 ommatidium were counted and the number of photoreceptors showing degeneration was scored. This was used to calculate the % of photoreceptors showing degeneration. This revealed 35.2%, 31.1% and 31.4% of photoreceptors were degenerated in DsRed2, 8B and 64 lines respectively. For the 38 repeat line 64.6% photoreceptors were degenerated again demonstrating enhanced age dependent cell loss because of 38 repeat expression (figure 4.5 C). In summary this data reveals expression of the 38 repeat construct causes a significant loss of neurons compared to controls at the same time point as measured by the number of remaining photoreceptor neurons at day 35.

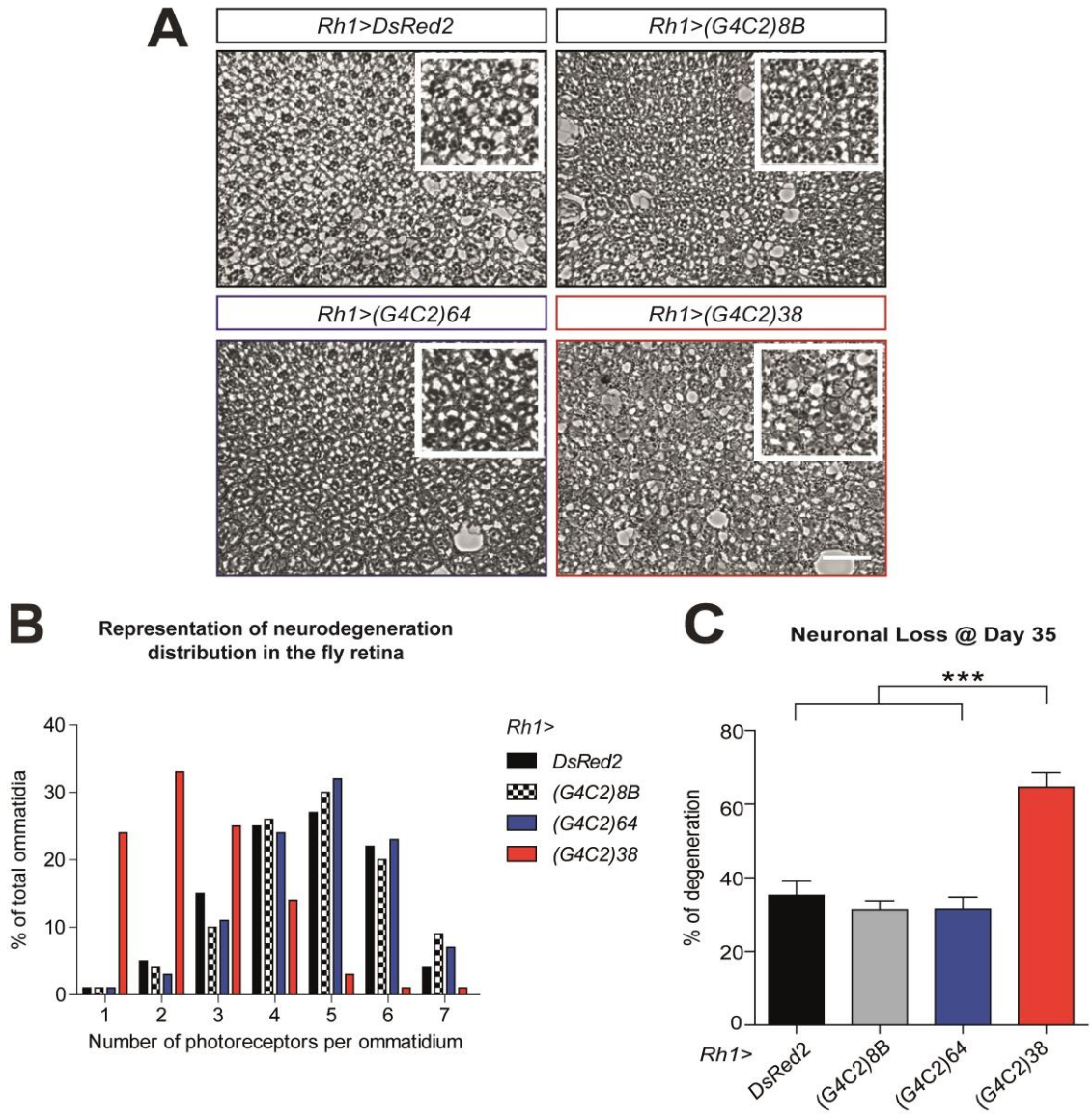

Figure 4.5 | Advanced neuron loss is seen in flies expressing 38 G_4C_2 repeats. **(A)** Semi-thin retinal sections of 35 days old flies over-expressing *DsRed2* (n=5), *(G_4C_2)8B* (n=6), *(G_4C_2)38* (n=6) or *(G_4C_2)64* (n=5), under the control of Rhodopsin1 promotor (*Rh1>*). Higher magnification of the retina structure is displayed on the top right inset. Scale bar, 20µm. **(B)** Histogram showing the distribution of photoreceptors in the ommatidia (displaying 1 to 7 photoreceptors per ommatidium) the distribution is shifted toward less photoreceptors per ommatidium for the 38 repeat line compared to the controls and the 64 repeat line ($p<0.0001$). A chi-square test on contingency table was performed, reported in appendix 10, table A10 **(C)** Histogram indicating percentage photoreceptor loss. A one-way ANOVA was performed which demonstrated a significant effect of genotype on photoreceptor loss ($F=23.50$, $p<0.0001$). Post-hoc analysis was performed using Bonferroni-Holm corrected multiple comparisons which can be found in appendix 10, table A10. Post-hoc analysis revealed the 38 repeat expressing flies show enhanced age related neurodegeneration compared to controls and the 64 repeat line; ***$p<0.0001$ with mean and SEM shown (n=5-6).

4.5 Discussion

4.5.1 Poly-GR expression correlates severe early onset motor impairment and neurodegeneration whereas poly-GP and poly-GA correlate with late onset motor impairment

This study has investigated a *Drosophila* model of C9ALS/FTD in which different lengths of uninterrupted G_4C_2 repeats were expressed in the 3'UTR of the disease-unrelated *DsRed2* transgene. Sense DPRs were detected in flies expressing ≥32 repeats; the levels and combinations of which correlated with the linker sequence 5' of the repeats. High levels of poly-GP and poly-GA were observed in the 32 and 64 repeat lines, which have the same 5' linker sequence, however much lower levels were seen in 38 and 56 repeat lines, both of which have their own unique 5' linker sequence. Hence for G_4C_2 repeats of a different length, the same linker produced similar DPR expression profiles, in agreement with findings that the sequence 5'-prime of repeats impacts RAN translation (Green et al., 2016). Of note, poly-GR was only detected in flies expressing 38 repeats.

The differential expression profiles thus facilitated assessment of the toxicity of different combinations and levels of RAN translated DPRs. This revealed high levels of poly-GA and poly-GP accumulation correlated with late onset age-related phenotypes, whereas poly-GR accumulation – only detected in the 38 repeat line – correlated with a rapid onset and progression of disease as measured by deficits in motor performance and neurodegeneration. This phenotype was not caused by the enhanced G_4C_2 RNA levels or the increased expression of DsRed2 seen in the 38 repeat line. This was confirmed by generating flies harbouring two copies of the 64 repeat construct which have significantly higher levels of DsRed2 protein compared to both flies with a single copy of the 64 repeat construct and the 38 repeat line. Day 5 climbing performance was not significantly different between the 2x 64 flies and the 1 x 64 flies, however were both significantly impaired compared to the 38 repeat line (appendix 2; figure A2). Taken together, these data indicate poly-GR and/or a combination of poly-GP/GP/GR (and not DsRed2 expression or G_4C_2 RNA levels) causes the early-day 5 behavioural phenotype seen in the 38 G_4C_2 repeat line.

Poly-GP and poly-GA at high levels has no impact on the motor performance of flies at day 5 and 20; only when flies are aged to day 40 is a mild reduction in climbing performance and motor activity seen. This motor impatient is unlikely to be attributable to the effects of the G_4C_2 RNA as no motor deficits are seen at day 40 in the low poly-GP and poly-GA expressing 56 repeat flies; which not only have a similar repeat length to the 64 but also have the same transcript levels based on DsRed2 expression (appendix 2; figure A1). Therefore, this data suggests this late-onset phenotype is caused by poly-GP and/or poly-GA. The different behavioural phenotypes seen and their relation to the DPR expression profile and levels is summarised in table 2.

Table 2 | The DPR expression profile in relation to the motor phenotypes observed. Constructs with the same linker sequence produce the same combination and levels of DPRs Severe early onset phenotypes correlate with poly-GR expression; high levels of poly-GP and poly-GA expression correlate with late onset motor phenotypes.

# G4C2 repeats	Expression Level	DPR Type	DPR Levels	Motor Phenotype
≤ 8	n/a	none	n/a	not at day 40
56	+	GA/GP	+/++	not at day 40
32	+	GA/GP	+++/+++	at day 40
64	+	GA/GP	+++/+++	at day 40
38	++	GA/GP/GR	++/+/+	at day 5

*Expression level based on quantitative western blotting of DsRed2 protein levels; DPR levels based on quantification of DPR levels by western blotting

Which of the DPR species is responsible for the severe early onset phenotype seen in the 38 repeat line? The individual DPRs can be studied in isolation of the potentially pathogenic G_4C_2 RNA and other DPRs. This is done by making recombinant DPR proteins and using the degeneracy in the genetic code. From these models there is a consistent finding that the arginine DPRs – poly-GR and poly-PR – are high neurotoxic (Freibaum and Taylor, 2017). Poly-GR and poly-PR are particularly detrimental when expressed in flies – with all studies using *Drosophila* to model their toxicity showing they cause severe neurodegeneration and a marked reduction in their survival and lifespan (Mizielinska et al., 2014; Wen et al., 2014; Freibaum et al., 2015; Yang et al., 2015; Boeynaems et al., 2016; Lee et al., 2016). RNA only models have shown the G_4C_2 RNA is not toxic to flies (Mizielinska

et al., 2014; Tran et al., 2015). The arginine rich DPRs have recently been shown to bind to proteins with low complexity domains (LCDs) – many of these proteins form parts of membraneless organelles such stress granules, nuclear speckles and nucleoli - impairing their assembly, dynamics, and function (Lee et al., 2016, Lin et al 2016., Boeynaems et al., 2017). Other pathogenic mechanisms associated with poly-GR and poly-PR toxicity include disrupted nucleocytoplasmic transport (Jovičić et al., 2015; Boeynaems et al., 2016a), disrupted mitochondrial function (Lopez-Gonzalez et al., 2016), nucleolar stress (Tao et al., 2015) and a blocking of global protein translation (Kanekura et al., 2016).

It is also possible a combination of poly-GP/GA/GR cause toxicity in the 38 repeat flies. Mori et al. (2016) found TDP-43 mislocalisation was enhanced in cells expressing both poly-GA and poly-GR. Further in the human brain there is a poor correlation between individual DPRs and RNA foci with regional neurodegeneration. Vatsavayai et al. (2016) suggest a combination of the different features in a neuron may be toxic; in support of this the authors found in two C9FTD cases that 27% of medial pulvinar thalamus (mPULV) neurons contained a combination of C9 pathologies and was one of the most degenerated regions in both cases. However, it has also been shown poly-GA and poly-GR when co-expressed is neuroprotective in flies; as poly-GA sequesters poly-GR into its inclusions thereby mitigating its toxicity (Yang et al., 2015). Further the present study has shown poly-GP and poly-GA co-expression at higher levels together produces a late-onset phenotype. Hence based on the above it is most probable that the severe early behavioural phenotype and neurodegeneration seen in the 38 repeat line is caused by the poly-GR peptide.

A late-onset day 40 phenotype is seen in the 32 and 64 repeat flies expressing high levels of poly-GP and poly-GA. This phenotype is likely caused by the high levels of poly-GP and poly-GA seen in these two lines, as the 56 repeat line which produces much lower levels of poly-GP and poly-GA but shows no phenotype at day 40. This is despite the 56 line having a repeat length like the 64 and similar transcript levels based on DsRed2 expression. This therefore indicates the G_4C_2 RNA is not behind this late onset phenotype. Hence, this data suggests poly-GP and poly-GA cause a late disease onset, and the expression levels of these DPRs appears to be an important factor in the timing of this. Further the expression pattern and levels of poly-GP in the 32 and 64 repeat flies is unchanged at days 5, 20 and 40 (appendix 4;

figure A4 A and A4 B). The finding that levels of poly-GP is still the same at different ages suggests poly-GR is produced at the later day 40 timepoint at which the motor impairment is observed. This model is unable to distinguish whether the combination of both DPRs or just one of poly-GP and poly-GA is the toxic species behind the late onset phenotype. Zu et al. (2013) using a G_4C_2 construct to overexpress poly-GP (via an insertion of an ATG in the poly-GP frame) did find increased cell death. However, in protein only models studying poly-GP toxicity (May et al., 2014., Wen et al., 2014., Freibaum et al., 2015., Tao et al., 2015, Yamakawa et al., 2015., Lee et al., 2016; Lee et al., 2017) only one of these studies have identified a very mild toxic role for poly-GP through impairments of the ubiquitin-proteasome system (Yamakawa et al., 2015). Poly-GP alternative codon *Drosophila* have been generated (Freibaum et al., 2015, Lee et al., 2016) and although the peptide did not cause any obvious external eye degeneration, these flies were not aged. Poly-GP could exert some toxicity after accumulating overtime. Indeed, no studies have investigated whether exposure to poly-GP alone causes any age-related deficits in motor performance or reduced life-span. In addition, cerebellar poly-GP has been associated with cognitive deficits in expansion carriers (Gendron et al., 2015). In the 56 repeat flies poly-GP was nuclear, forming paranucleolar accumulations as has been previously reported by in human C9 frontal cortex (Schludi et a., 2015). In the higher expressing 64 repeat flies poly-GP showed a predominantly diffuse cytosolic localisation. Diffuse cytosolic poly-GP may have a role in toxicity; whereas nuclear poly-GP appears to not confer toxicity. The precise subcellular localisation of the DPRs may alter their toxicity, for example it has been shown rerouting poly-GA aggregates from the cytoplasm to the nucleus prevents cytosolic TDP-43 accumulation (although the toxicity was not assessed) and the nuclear localisation of poly-GR and poly-PR is necessary for their toxicity (Rudich et al., 2017).

Like poly-GP, when poly-GA is expressed in isolation in flies no early-onset neurodegeneration is seen (Freibaum et al., 2015; Lee et al., 2016). Poly-GA on the other hand has been shown to be at least mildly toxic in several different model systems, although much less toxic than the arginine DPRs when directly compared in the same model system (Freibaum and Taylor., 2017). Protein only poly-GA murine models have shown the peptide causes age related impairments in cognition and motor performance in addition to inflammation and neurodegeneration (Zhang et

al., 2016; Schludi et al., 2017). Whilst no poly-GA external eye degeneration is observed, such flies do show a reduced lifespan compared with controls when aged (Mizielinska et al., 2014). Although other fly studies have reported poly-GA is neuroprotective (Yang et al., 2015). Poly-GA in the 64 repeat flies formed paranucleolar accumulations reminiscent of the nuclear inclusions reported in C9 human frontal cortices (Schludi et al., 2015); further paranucleolar poly-GA aggregated with heterochromatin and H3K9me2 (a marker of transcriptional repression) in the human frontal cortex, suggesting a link between poly-GA, gene transcription and transcriptional silencing (Schludi et al., 2015).

4.5.2 Conclusion

In summary the different DPR profile and levels of expression observed in the G_4C_2 flies used in this study provided an opportunity to compare the behavioural toxicity caused by different combinations and levels of DPRs RAN translated from the G_4C_2 expansion. Severe early onset deficits in climbing performance and motor activity were only seen in the 38 repeat line - the only line in which poly-GR could be detected. Late onset behavioural phenotypes were seen in aged 32 and 64 repeat flies expressing high levels of poly-GP and poly-GA compared to controls and importantly the 56 repeat line which produced significantly lower levels of both poly-GP and poly-GA. The 56 repeat flies had similar transgene levels to the 32 and 64 repeat flies suggesting G_4C_2 RNA is not contributing to the age related phenotype seen in these flies. A better understanding of the localisation and behaviour of poly-GP and poly-GA in aged flies may potentially provide insight into the age-related motor impairment seen in the 32 and 64 repeat flies. Hence severe early behavioural phenotypes correlated with the presence of poly-GR whereas poly-GP and/or poly-GA cause late-onset behavioural impairment dependent on their expression level.

Chapter 5: Generation of alternative codon DPR only constructs and characterisation of the toxicity and localisation of alternative codon ATG derived poly-GR and poly-GA

The expression of poly-GR correlates with an early-onset severe motor impairment and neurodegeneration in flies; whereas high levels of poly-GP and poly-GA correlate with an age-dependent decline in motor performance. To study DPR toxicity further alternative codon constructs, designed to produce a DPR (e.g. poly-GR) from non-G_4C_2 codons and not from RAN translation were generated. Such constructs allow individual DPRs to be investigated independently of each other and allow DPR toxicity to be studied in isolation of G_4C_2 RNA toxicity. Alternative codon constructs producing different lengths of pure DPR (8 and 64) were designed for poly-GA, poly-GR, poly-PR and poly-PA. A poly-GP alternative codon construct was unable to be synthesised due to the high GC content of the codons required despite avoiding using the G_4C_2 sequence. The previous chapters focused on RAN translation and toxicity of the sense DPRs as the anti-sense DPRs were either undetectable or not produced in the DsRed2-G_4C_2 repeat flies studied in this work. Hence characterisation and study of the DPRs focused on the poly-GA and poly-GR constructs for the remainder of this thesis.

5.1 Alternative codon constructs used to produce DPRs from non G_4C_2 RNA

Alternative codon constructs producing different lengths of pure DPR protein were designed for all DPRs except poly-GP due to failures in generating and cloning a sequence with such a high GC content (despite the use of non G_4C_2 codons) (figure 5.1). The constructs included an insect-specific Kozak sequence (xACAAA) (Schleich et al., 2014) for efficient translation as well as 5' HA and 3' Myc tags for detection *in vivo*. Care was taken to avoid the presence of any G4C2 repeats in the design of the constructs or any repetitive stretches of the same codons. These constructs have a traditional ATG codon hence the individual DPRs are translated from normal canonical mechanisms rather than RAN translation. Both short length 8 DPR repeat and longer length 64 DPR repeats constructs were generated (figure 5.1). These protein only constructs were sent to BestGene, USA for transgenisis. All constructs were inserted into the genome at a common insertion site via φC31-mediated site-specific integration into flies harbouring an *attP* insertion at site attp40 on the second chromosome (y^1 w^{67c23}; P{CaryP}attP40). Genomic PCR confirmation of insertion onto the attp40 landing site was confirmed by BestGene, USA. As the previous chapters investigating DPR toxicity in the DsRed2-G_4C_2 repeat flies focused on the sense DPRs rather than the anti-sense- the reminder of this chapter focuses on the characterisation of the poly-GA and poly-GR alternative codon constructs. Full length sequences of the DPR constructs generated in this study can be found in appendix 5.

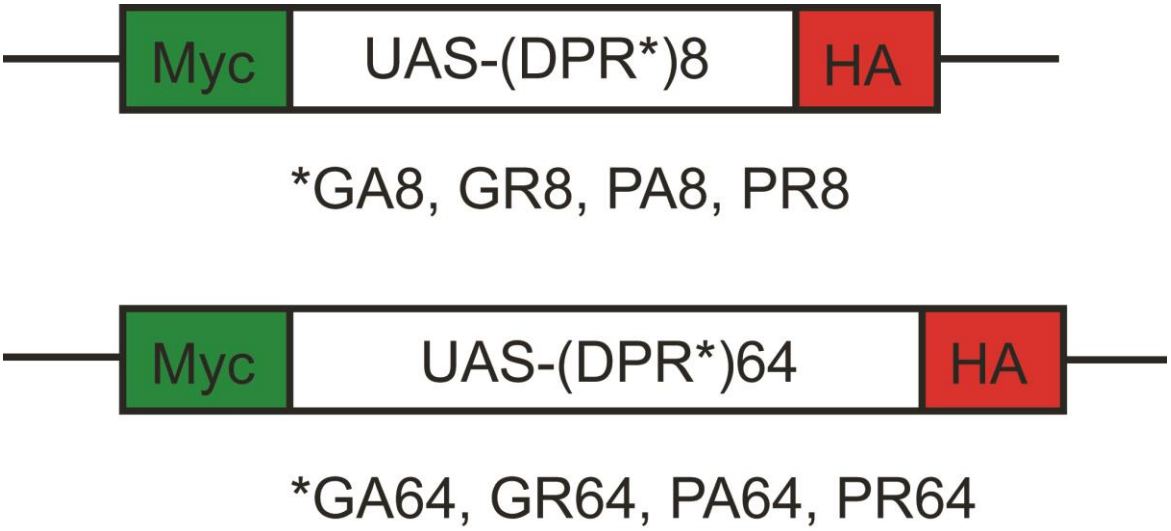

Figure 5.1 | Schematic of the alternative codon, non G_4C_2-derived DPR constructs used in this study. Schematic depiction of non-G4C2 alternative codon constructs designed to produce different lengths of pure DPR (8 DPR repeats and 64 DPR repeats) for poly-GA, poly-GR, poly-PR and poly-PA, flanked by 5' Myc tag and 3' HA tag.

5.2 Poly-GR cause severe neurodegeneration when expressed in the Drosophila eye

The arginine-rich DPRs poly-GR and poly-PR have been demonstrated to be highly toxic in several studies; particularly in *Drosophila* models (Mizielinska et al., 2014; Wen et al., 2014; Yang et al., 2015; Freibaum et al., 2015; Boeynaems et al., 2015; Lee et al., 2016). To assess the toxicity of the poly-GR64 alternative codon DPR-only constructs designed in this study, the poly-GR8, poly-GR64 and poly-GA64 constructs were expressed in the fly eye using the *GMR-GAL4* driver (figure 5.2). As reported previously in the literature, the poly-GR64 constructs show a severe degenerative phenotype in eye (figure 5.2). No change in external eye morphology is seen in flies expressing poly-GR8 or poly-GA64 (figure 5.2), These data replicates what has been shown in several *Drosophila* studies of DPR toxicity, that poly-GR toxicity is length dependent and poly-GA does not cause external degeneration when expressed in the fly eye (Mizielinska et al., 2014; Wen et al., 2014; Yang et al., 2015; Freibaum et al., 2015; Boeynaems et al., 2016; Lee et al., 2016).

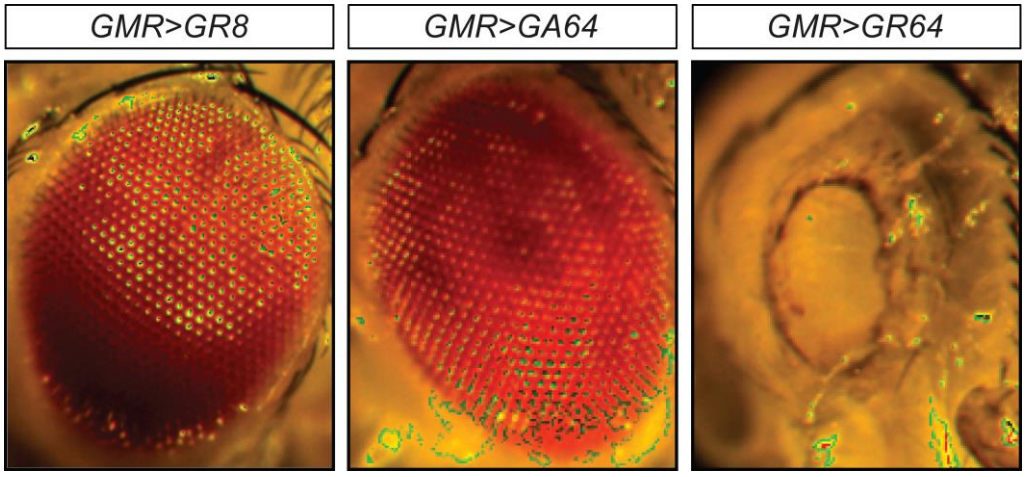

Figure 5.2 | The poly-GR64 construct when expressed in the *Drosophila* compound eye causes severe degeneration. The poly-GR8, poly-GA64 and poly-GR64 constructs were expressed in the compound eye of the fly using the *GMR-GAL4* driver. Poly-GR8 and poly-GA64 expression did not alter the external morphology of the fly eye. Poly-GR64 on the other hand caused a severe degeneration of the eye.

5.3 Poly-GA and poly-GR show different subcellular localisations

In vitro models of epitope and GFP tagged DPRs produced from protein only constructs have shown different and distinct subcellular localisation and aggregation properties of the 5 DPRs (May et al., 2014; Zhang et al., 2014; Wen et al., 2014 Yamakawa et al., 2015; Tao et al., 2015). To confirm the successful expression of the individual DPRs from the poly-GA64 and poly-GR64 constructs the *UAS-DPR* transgenes were expressed in the larval eye disc using the *GMR-GAL4* driver (figure 5.3). L3 larval eye discs were dissected and stained with HA antibody. Poly-GA64 (*GMR>GA64*) localisation in the photoreceptor neurons of the L3 larval eye disc is predominantly diffusely nuclear and cytoplasmic, staining is seen both in the nucleus of the ommatidium and the surrounding cytosol, although small compact spherical inclusions are also observed. Poly-GR64 (*GMR>GR64*) in the L3 larval eye disic is exclusively cytoplasmic in photoreceptor neurons, some perinuclear staining is also seen with the poly-GR surrounding the nuclei of individual photoreceptors of a single ommaditium. Although these DPRs form inclusions in the patient brain (Schludi et al., 2015, Mackenzie 2015) only poly-GA formed aggregates at this early time point consistent with previous studies and its compact aggregate structure (May et al., 2014; Chang et al., 2016; Freibaum and Taylor, 2017).

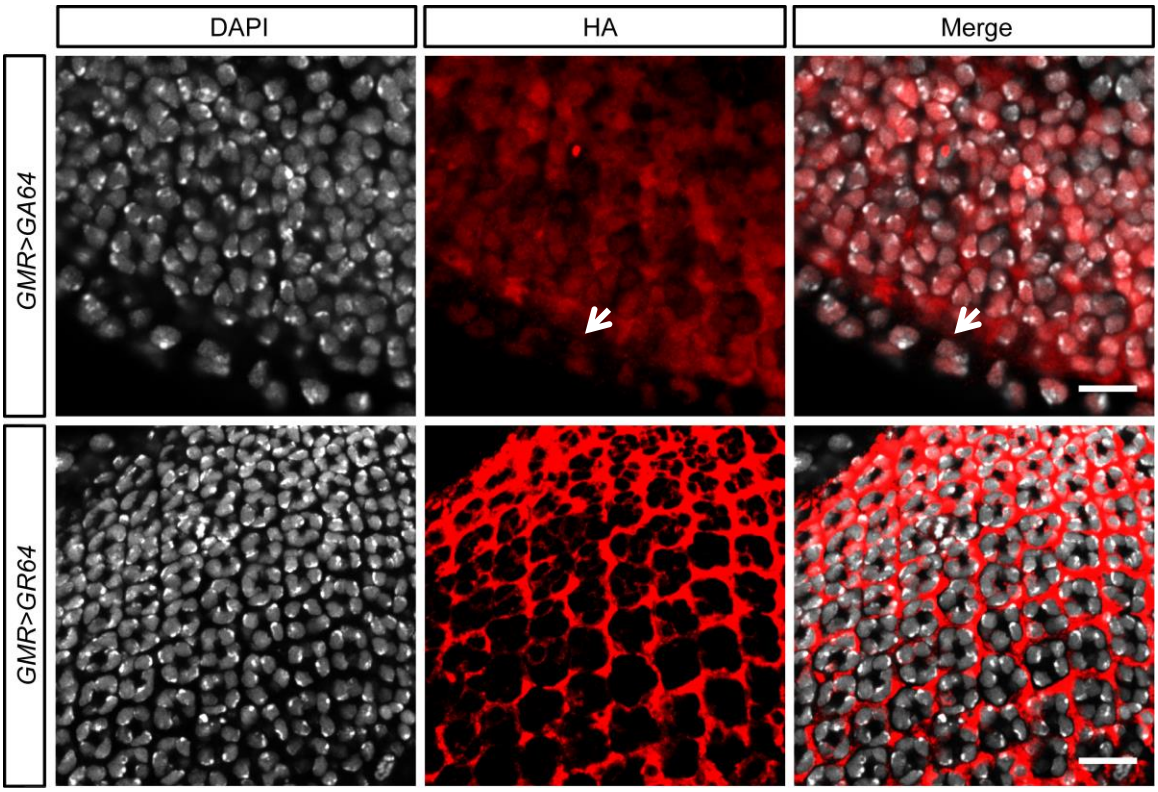

Figure 5.3 | The DPRs show diffuse subcellular localisations with only poly-GA forming inclusions in L3 larval eye discs. Confocal images of L3 larval eye discs immunolabelled for anti-HA which reveals DPR accumulation. Poly-GA is diffusely nuclear and cytoplasmic although forms compact inclusions at this early time point (white arrow); Poly-GR at this stage is predominantly diffusely cytoplasmic although some pernuclear staining is also seen. Scale bars, 10μm.

5.4 Poly-GR localisation changes from cytoplasmic to nuclear over time

No cytoplasmic or nuclear aggregates of poly-GR were seen in the L3 larval eye disc. Previous studies have shown the arginine containing DPRs poly-GR and poly-PR localise to the nucleus and form nuclear inclusions in human post mortem tissue in both *in vivo* and *in vitro* models of DPR toxicity (May et al., 2014; Wen et al., 2014; Boeynaems et al., 2016). Protein translation occurs in the cytoplasm hence poly-GR and poly-PR must actively translocate into the nucleus. To study the dynamics of the nuclear aggregation of poly-GR in *Drosophila*, neuronal tissue was probed at two different time points. Poly-GR64 localisation was studied in Drosophila photoreceptors using the *GMR-GAL4* driver. Eye discs were dissected at two stages – the L3 larval stage and ~48 hours later at the pupal stage (figure 5.4A). Poly-GR initially shows a diffuse cytoplasmic distribution in L3 larval eye discs, however rapidly within 48 hours, at the pupal stage, poly-GR64 localisation dramatically changes, where the protein is found predominantly in the nucleus and also perinuclear (figure 5.4A). Furthermore, nuclear poly-GR64 formed distinct nuclear inclusions (figure 5.4A) Hence the poly-GR64 peptide rapidly translocated from the cytoplasm into the nucleus where it aggregates.

To circumvent the pupal lethality of the poly-GR64 construct using the pan neuronal driver Elav^{C155} driver (data not shown) the *EB1-GAL4* driver was utilised to study poly-GR localisation in adult brain. *EB1-GAL4* specifically targets the *Drosophila* upper motor neurons and becomes active in the latter pupal stages of development and stays activated during the lifespan of the fly in the ellipsoid body ring upper motor neurons (Diaper et al., 2013a) (figure 5.4B). Previous work overexpressing the *Drosophila* homologue of TDP-43, *TBPH*, has shown the loss of this small population of neurons does not alter the lifespan of flies unlike pan neuronal expression of *TBPH*, therefore this driver is ideal to study poly-GR at a cellular resolution (Diaper et al., 2013a). The *EB1-GAL4* driver is co-expressed with a membrane-bound form of GFP (*UAS-mCD8::GFP*) which allows axons and synaptic terminals to be seen. *EB1-GAL4* driven poly-GR64 and control brains were dissected and stained with anti-HA. EB1 neurons were identified by membrane bound GFP auto-fluorescence. Adult brains were dissected at two time points – day 2 and day 18 (figure 5.4C). Day

18 was chosen as the end time-point as flies expressing poly-GR64 in EB1 neurons displayed no obvious motor impairment at this age, hence it was decided to dissect the flies rather than age them further. No HA signal was detected in EB1 neurons in the control (*EB1>mCD8::GFP*) flies (figure 5.4C). For EB1 driven poly-GR64 as was the case for the L3 eye discs, poly-GR at day 2 is diffuse and cytoplasmic however in aged day 18 fly brains the protein is perinuclear and localised to the nucleus where inclusions are observed (figure 5.4C). *EB1-GAL4* dissections were performed by KCL BSc student Rachel Hall under my supervision.

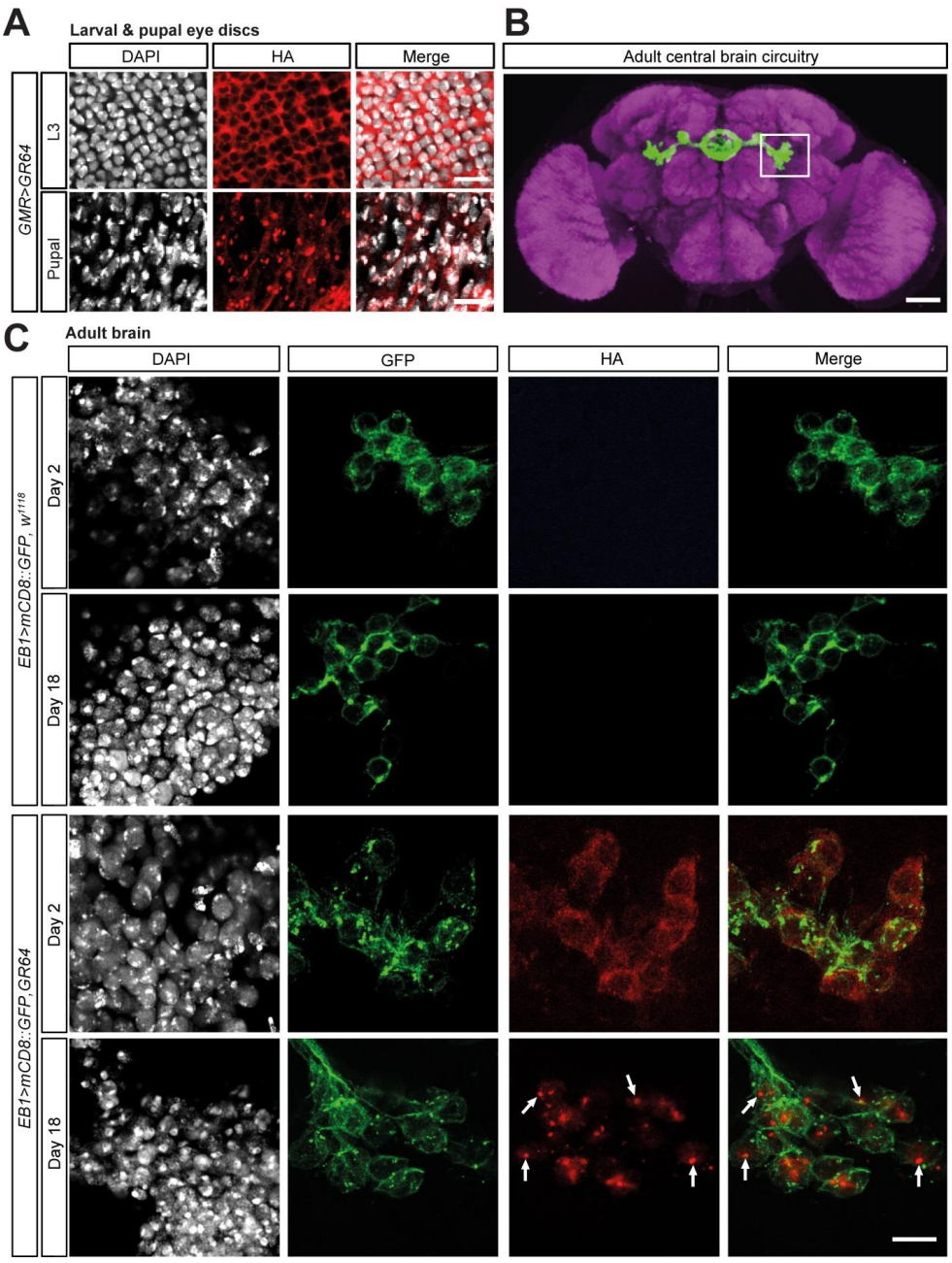

Figure 5.4 | Poly-GR accumulation changes over time; from diffusely cytoplasmic to the formation of predominantly nuclear aggregates. (A) Confocal images of L3 and pupal eye discs of *GMR>GR64* flies immunolabeled for HA and DAPI, indicating poly-GR distribution. Localisation of HA/poly-GR initially reveals diffuse cytoplasmic distribution and within 48 hours (L3 to pupal transition) also shows nuclear inclusions. **(B)** Confocal image of *EB1>mCD8::GFP* adult *Drosophila* brain immunolabeled with anti-Brp (magenta) and anti-GFP (green) highlighting *Gal4/UAS* activity targeted to ellipsoid body circuitry (from Zoe Ludlow PhD thesis; KCL, Hirth lab). **(C)** Confocal images of ellipsoid body neurons at day 2 and day 18 expressing *UAS-mCD8::GFP* either alone or in conjunction with *UAS-GR64*. Immunolabeling for HA reveals diffuse cytoplasmic and perinuclear distribution of poly-GR by day 2 and by day 18 also identifies nuclear, perinuclear and cytoplasmic inclusions (arrows). Scale bars, 10µm.

5.5 Poly-GR nuclear inclusions do not co-localise with the nucleolar marker fibrillarin

The preceding section has shown poly-GR rapidly forms nuclear inclusions *in vivo*. Previous research with both poly-GR and poly-PR alternative codon constructs in cellular models has demonstrated these nuclear aggregates frequently co-localise with the nucleolus (Schludi et al., 2015; Tao et al., 2016; Callister et al., 2016). To determine whether the nuclear inclusions of poly-GR64 co-localised with the nucleolus *Drosophila* L3 salivary glands and brains expressing poly-GR64 were co-stained with the nucleolar marker fibrillarin, a component of a nucleolar small nuclear ribonucleoprotein (figure 5.5A). Alternative codon poly-GR64 was expressed in the salivary gland cells of larvae using the *FKH-GAL4* driver. Salivary gland cells stained with HA did not detect any signal for the short repeat poly-GR8 construct (*FKH>GR8*). Poly-GA64 formed both cytoplasmic and nuclear aggregates; poly-GA64 nuclear inclusions did not co-localise with fibrillarin (figure 5.5A). Poly-GR64 formed nuclear aggregates, however these did not co-localise with the nucleolus (figure 5.5A). No co-localization was seen between nuclear poly-GR64 inclusions and nucleolar fibrillarin in salivary gland cells (figure 5,5A).

To confirm the findings in the salivary gland cells the poly-GR64 transgene was expressed in neurons using the pan neuronal $Elav^{C155}$ driver (figure 5.5B). L3 larval brains were dissected and stained with anti-HA and anti-fibrillarin. Poly-GR64 nuclear inclusions in the ventral nerve cord also did not co-localise with fibrillarin - although were adjacent to the nucleolus (figure 5.5B). Interestingly these findings mirror what is seem in patient tissue where 78% of nuclear poly-GR intranuclear inclusions are attached to the nucleoli - but never co-localising with the nucleolus (Schludi et al., 2015). Furthermore, double immunostaining of p62 and nucleolin revealed no colocalization (Schludi et al., 2015). Similar reports of a ‘nucleolar studding’ pattern for DPRs around but not directly within the nucleolus was also reported by Vatsavayai et al. (2016). It remains unclear how these inclusions interact with the nucleolar structure and whether this interaction leads to nucleolar dysfunction. Hence the intranuclear inclusions seen in the poly-GR64 resembles patient post-mortem poly-GR pathology unlike cellular protein only models of poly-

GR (Schludi et al., 2015; Tao et al., 2016; Callister et al., 2016; Vatsavayai et al., 2016).

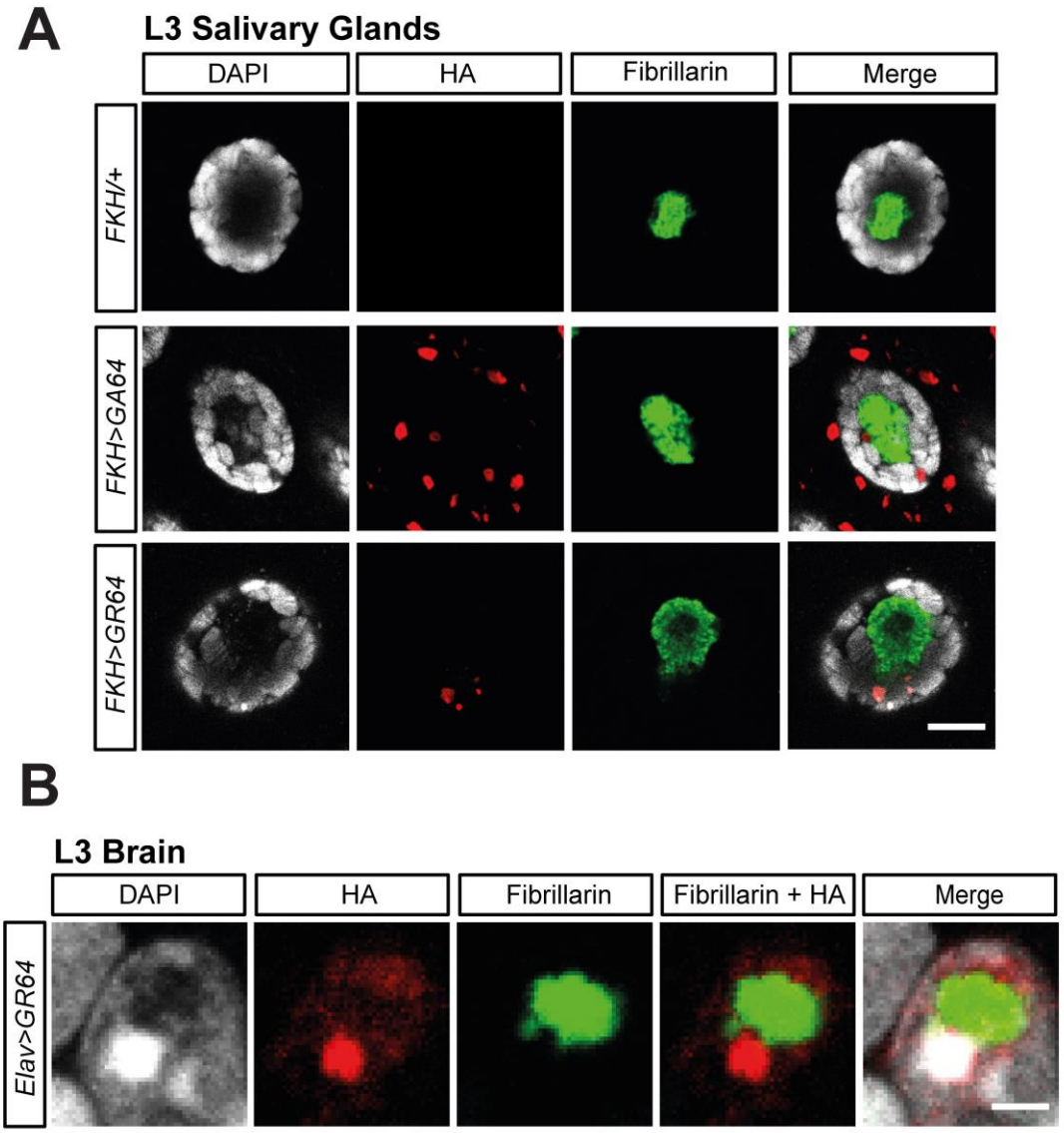

Figure 5.5 | Poly-GR forms nuclear inclusions that aggregate adjacent to the nucleolus. (A & B) Confocal images of L3 larval eye discs and larval ventral nerve cord immunolabeled for anti-HA and the nucleolar marker fibrillarin. Staining shows poly-GR nuclear inclusions that are adjacent to, but do not co-localise with the nucleolus. Scale bars, 10µm **(A)**, 5µm **(B).** .

5.6 Poly-GA but not poly-GR aggregates are p62 positive

The DPR inclusions in the patient brain are negative for TDP-43 but positive for p62 and ubiquitin (Mackenzie et al., 2014). p62 (or Sequestome-1) targets specific cargoes for autophagy (Rusten and Stenmark, 2010). Only overexpressed poly-GA from alternative codon constructs co-localises with p62 - none of the other DPRs, including nuclear poly-GR and poly-PR, formed p62 positive inclusions (May et al., 2014; Zhang et al., 2014; Schludi et al., 2015; Callister et al., 2016). *Drosophila* have one homologue of p62, *Ref(2)P* (Nezis et al,, 2008). *Ref(2)P* localises to age-related proteinaceous aggregates and aggregations induced by proteasome and autophagy inhibition and is necessary for the formation of aggregates in the fly brain (Nezis et al,, 2008).

Alternative codon poly-GR8, poly-GA64 and poly-GR64 was expressed in the salivary gland cells of larvae using the *FKH-GAL4* driver. The 288-G_4C_2 RNA only construct which does not produce DPRs was also included as DPR negative control. Salivary gland cells stained with an antibody specific for *Ref(2)P* revealed the protein to be predominantly cytoplasmic with a very weak nuclear stain in controls (*FKH/+, FKH>GR8*) (figure 5.6). In salivary gland cells expressing the 288-G_4C_2 RNA only construct (*FKH>288 RO*) *Ref(2)P* resembled the diffuse cytosolic staining seen in controls (figure 5.6). However, for poly-GA64 expressing salivary glands virtually all of the diffuse cytoplasmic *Ref(2)P* was gone – the protein was sequestered into poly-GA64 inclusions (figure 5.6). Interestingly for poly-GA64 inclusions with halo-like structures, *Ref(2)P* was observed in the centre of these accumulations – similar observations have been seen in both cellular models and the post-mortem brain for poly-GA and p62 (Schludi et al., 2015). Nuclear poly-GR64 inclusions on the other hand were not *Ref(2)P* positive, with the protein keeping its predominantly cytoplasmic localisation and weak nuclear stain (figure 5.6). This is unlike what is seen in patient tissue where poly-GR p62-negative inclusions only occur extremely rarely in the cytoplasm and nucleus; the vast majority of cytosolic and nuclear poly-GR inclusions are p62 positive (Schludi et al., 2015). Hence, these findings indicate that like current cellular protein only DPR models, *in vivo* poly-GR models do not fully replicate aggregates found in patient tissue.

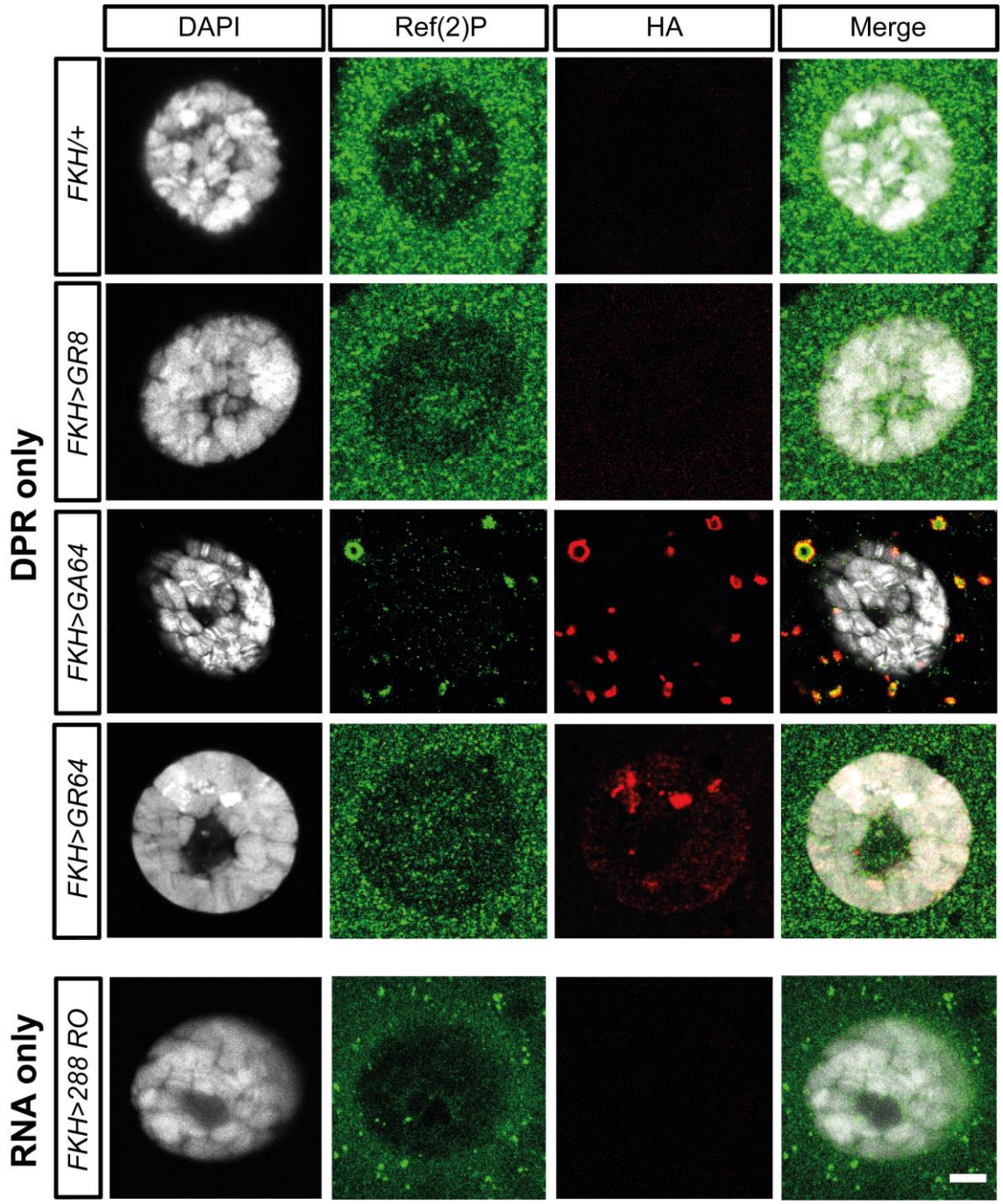

Figure 5.6 | Poly-GA, but not poly-GR inclusions are *Ref(2)P* positive. Salivary gland cells expressing poly-GA64, poly-GR64 and controls were stained with *Ref(2)P*, the *Drosophila* homologue of p62. *Ref(2)P* displayed a predominantly diffuse cytoplasmic localisation in controls (*FKH/+* and *FKH>GR8*) and the G_4C_2 RNA only line (FKH>288 RO). However this diffuse cytosolic *Ref(2)P* labelling is not present following poly-GA64 expression (*FKH>GA64*), with the protein being sequestered into poly-GA inclusions as in the human brain. Nuclear poly-GR inclusions are however not *Ref(2)P* positive. Scale bars, 10µm.

5.7 Discussion

5.7.1 Poly-GR when expressed in the *Drosophila* eye causes severe neurodegeneration

Poly-GR expression the fly eye caused severe external compound eye degeneration. This data recapitulates the findings from previous fly models studying DPR toxicity demonstrating the arginine DPRs are the main drivers of toxicity in C9ALS/FTD (Mizielinska et al., 2014; Wen et al., 2014; Yang et al., 2015; Freibaum et al., 2015 Boeynaems et al., 2015; Lee et al., 2016). Further this also strongly suggests poly-GR is most likely responsible for the enhanced motor phenotype and retinal degeneration seen in the 38 G_4C_2 repeat line characterised in this study.

5.7.2 Poly-GR localisation changes from cytoplasmic to nuclear over time but inclusions do not co-localise with the nucleolus

Poly-GR64 stained for at early time points (the L3 larval eye disc and day 2 adult EB1 neurons) show a diffuse cytoplasmic labelling; however, as flies are aged the protein translocated into the nucleus where it formed inclusions. This transition into the nucleus is likely down to the highly negative charge of the peptide (Freibaum and Taylor., 2017) compared to the positively charged histones that package and order DNA into nucleosomes. A nuclear accumulation of the arginine DPRs has been observed cellular models of DPR toxicity (Zhang et al., 2014; May et al., 2014; Wen et al., 2014; Tao et al., 2015; Yamakawa et al., 2015; Callister et al., 2016; Khosravi et al., 2017; Lee et al., 2017). Poly-GR could be actively transported into the nucleus by mimicking a nuclear localisation signal (NLS). A non-classical group of NLSs that are arginine–glycine-rich, known as RG-NLSs, have been identified in yeast proteins including yeast hnRNP A1 (Cautain et al., 2014). Poly-GR has already been shown to interact with the LCD domain of hnRNAP A1 altering its biophysical properties (Lee et al., 2016). These RG-NLSs have similar characteristics to the non-classical PY-NLS that consists of an N-terminal hydrophobic motif, a central arginine and the C-terminal PY sequence. Proteins with the PY-NLS motif are imported into the nucleus via the transportin-1 receptor (Cautain et al., 2015). Transportin-1 is a strong enhancer and suppressor of poly-PR neurodegeneration in a *Drosophila* and yeast

screen (Jovičić et al., 2015; Boeynaems et al., 2016a). Boeynaems et al. (2017) performed computational docking simulations predicting that poly-PR interacts with transportin-1 suggesting the peptide could potentially compete with transportin-1 for NLS containing protein cargoes. In support of this the RNA-binding protein Elav, a known binding partner of transportin-1, accumulated in the cytoplasm of flies expressing poly-PR. Hence like poly-PR, poly-GR, based on its glycine and arginine residues, might also bind to transportin-1 and be transported into the nucleus where it accumulates and forms inclusions. It is interesting to note however that FUS is a known transportin-1 cargo; yet FUS pathology is not seen in C9ALS/FTD. Furthermore, both poly-GR and poly-PR interact with numerous nuclear proteins by binding to their LCDs; including RNA binding proteins, nucleolar proteins, nuclear speckle proteins, cajal body proteins (Lee et al., 2016). Additionally, poly-PR has been shown to bind to the FG-repeats within the central channel of the NPC (Shi et al., 2017).

Nuclear poly-GR64 inclusions were found to be adjacent to the nucleolus in *Drosophila* L3 larval salivary glands and neurons; although did not directly co-localise with the nucleolus as has been reported in cellular models (Zhang et al., 2014; May et al., 2014; Wen et al., 2014; Tao et al., 2015; Schludi et al., 2015; Callister et al., 2016). Several studies have reported nucleolar poly-GR cause's nucleolar enlargement and impaired ribosome biogenesis by binding to the LCD of nucleolar proteins perturbing phase transitions in the liquid compartment of this membraneless organelle (Kwon et al., 2014; Tao et al, 2015; Callister et al., 2016; Lee et al., 2016). In the human brain however, poly-GR inclusions do not localise to the nucleolus like in cellular models; rather they are para-nucleolar (Schludi et al., 2015). Schludi et al. (2015) observed no co-localisation between poly-GR and the nucleolar marker fibrillarin in frontal cortex neurons; leading the authors to question the role of nucleolar stress in pathogenesis as unlike cellular models in post-mortem tissue and the structure and size of the nucleolus is unaffected (Schludi et al., 2015). A more recent study by Mizielinska et al. (2017) investigating nucleolar dysfunction in C9FTD frontal cortex revealed nucleolar volume in neurons is reduced in patient brains compared to controls. Intriguingly although nucleolar volume was on average lower in the C9FTD frontal cortex – nucleolar volume was increased in neurons with cytoplasmic poly-GR inclusions. Increased nucleolar volume was also seen *in vivo* in

a *Drosophila* poly-GR protein only model, in this model (as in the poly-GR64 flies used in this study and what was reported by Schludi et al. (2015) previously in patient brains) poly-GR inclusions did not co-localise with fibrillarin but were adjacent to it. Mizielinska et al. (2017) demonstrate poly-GR does not need to directly aggregate with nucleolar proteins to cause nucleolar dysfunction; showing cytoplasmic poly-GR causes nucleolar dysfunction in patient brains and *in vivo*. To explain this Mizielinska et al. (2017) propose that in these neurons with a high enough poly-GR load that cytoplasmic or nuclear inclusions are present - that soluble poly-GR, undetectable by immunofluorescence is interacting with nucleolar proteins and disrupting nucleolar functions. Hence despite the fact poly-GR in patient brains and the poly-GR64 flies used in this study does not co-localise with nucleolus, the peptide still can cause nucleolar dysfunction.

5.7.3 Poly-GA expression has no obvious effect on external eye morphology but forms cytoplasmic and nuclear inclusions that are p62 positive

No changes in external eye morphology were observed for poly-GA. This is consistent with previous *Drosophila* models in which poly-GA repeats have no toxic effect when expressed in the eyes of young flies (Mizielinska et al., 2014; Wen et al., 2014; Yang et al., 2015; Freibaum et al., 2016). However it has been shown previously that poly-GA causes a reduction in lifespan from day 50 onwards when expressed pan-neuronally in *Drosophila* (Mizielinska et al., 2014). Furthermore poly-GA/poly-GP correlated with a late-onset motor impairment in the 32 and 64 G_4C_2 repeat lines used in this study. This suggests either poly-GA induced neuronal toxicity slowly accumulates over time before any overt behavioural toxicity is seen or the ageing nervous system is more vulnerable to poly-GA inclusions. Hence insight into poly-GA toxicity and evidence for poly-GA induced neurodegeneration in flies will come from a characterisation of neuronal poly-GA localisation and behaviour in flies expressing poly-GA at ages in which phenotypes are seen.

Poly-GA formed both cytoplasmic and nuclear inclusions which were *Ref(2)P* positive - *Ref(2)P* is the *Drosophila* homologue of p62. In the human brain DPR inclusions are p62 positive, hence the poly-GA inclusions produced from the poly-GA64 constructs resemble what is seen in the human post mortem brain (Schludi et

al., 2015). Nezis et al. (2008) demonstrated previously that *Ref(2)P* localises to age-induced protein aggregates in the fly brain and aggregates induced by deficits in autophagy or the proteasome. This suggests poly-GA could cause defective autophagy or proteasome dysfunction and such processes may contribute to the reduced lifespan observed in poly-GA flies by Mizielinska et al. (2014) and the late onset motor impairment in the 32 and 64 G_4C_2 repeat flies studied in this thesis.

5.7.4 Conclusion

In summary this chapter has characterised alternative codon ATG driven poly-GA and poly-GR expression in a *Drosophila* model generated during this project. Poly-GA forms cytoplasmic and nuclear *Ref(2)P* positive inclusions; despite this poly-GA did not cause external degeneration of the *Drosophila* eye replicating what has been shown in several *Drosophila* studies of DPR toxicity (Mizielinska et al., 2014; Wen et al., 2014; Yang et al., 2015; Freibaum et al., 2015). Poly-GR on the other hand formed nuclear inclusions adjacent to the nucleolus and was highly toxic; indicating poly-GR is the toxic species causing the severe early motor impairment in the 38 G_4C_2 repeat flies.

Chapter 6: TBPH localisation in flies expressing either poly-GA, poly-GR or G_4C_2 RNA only

In addition to DPR and RNA pathology, patients harbouring the G4C2 expansion are also characterised by mislocalisation and neuronal cytoplasmic inclusions of the RNA binding protein TDP-43 (Al-Sarraj et al., 2011). TDP-43 pathology, unlike DPR pathology, shows strong correlations with both clinical features and neurodegeneration (Mackenzie et al., 2014; Gomez-Deza et al., 2015,;Mackenzie et al., 2015). However, how TDP-43 becomes pathogenic in expansion carriers is currently unknown. Murine models show RNA foci and RAN dipeptides develop early, but are not sufficient to drive neurodegeneration in the absence of TDP-43 pathology (O'Rourke et al., 2015; Peters et al., 2015), however neurodegeneration is seen murine models that present TDP-43 pathology (Chew et al., 2015; Liu et al., 2016c). Interestingly in one of these models TDP-43 aggregation occurred at end-stage, after the formation of both DPRs and RNA foci (Liu et al., 2016c). Furthermore post-mortem analysis on patients with the *C9ORF72* expansion who have died early of the disease show abundant DPR pathology only; further indicating DPR pathology precedes that of TDP-43 (Proudfoot et al., 2014; Baborie et al., 2015; Vatsavayai et al., 2016). An amyloid cascade hypothesis for C9 mediated neurodegeneration has been proposed in which specific stressors causes TDP-43 aggregation and cytoplasmic mislocalisation (Edbauer and Haass, 2016). The above data suggests DPRs and/or RNA foci act as such stressors, causing TDP-43 dysfunction leading to subsequent neurodegeneration (Edbauer and Haass, 2016).

6.1 Cytoplasmic TBPH accumulation is a consequence of poly-GR expression

The *Drosophila* homologue of TDP-43 is known as *TAR DNA-binding protein-43 homolog* (*TBPH*). *TBPH* shares all the major protein domains found in human TDP-43 (Diaper et al., 2013a) and has been shown to have conserved functions in relation to RNA processing (Romano et al., 2014). Both neuronal overexpression and knockdown of *TBPH* is toxic to flies, causing motor impairment and age-related neurodegeneration (Diaper et al., 2013a). *TBPH* localisation was investigated in the salivary gland cells of controls; DPR only fly lines and the RNA only 288 G_4C_2 repeat fly line generated by Mizielinska et al. (2014).

The *FKH-GAL4* driver was used to express either poly-GR64 or poly-GA64 in salivary gland cells. This driver was chosen as the salivary gland cells have a large nuclei and cytoplasm that can easily be distinguished. In control salivary gland cells (*FKH/+* and *FKH>GR8*) *TBPH* labelling was predominantly nuclear, a weak cytoplasmic *TBPH* signal was also seen as has been previously reported in *Drosophila* neurons (Diaper et al., 2013a). *TBPH* localisation was indistinguishable from the controls for G_4C_2-288 repeat RNA only expressing salivary gland cells, remaining predominantly nuclear. *TBPH* was also mainly nuclear in salivary gland cells expressing poly-GA, however cytosolic *TBPH* co-localised with poly-GA64 inclusions in the cytoplasm of salivary gland cells. Poly-GR64 expression in salivary gland cells caused a dramatic increase in cytoplasmic *TBPH* – without a noticeable nuclear loss. A cytoplasmic accumulation of *TBPH* without noticeable nuclear loss was also seen in when poly-GR64 was expressed in photoreceptor neurons but not controls (*GMR/+* and *GMR>GR8*) (appendix 6; figure A6).

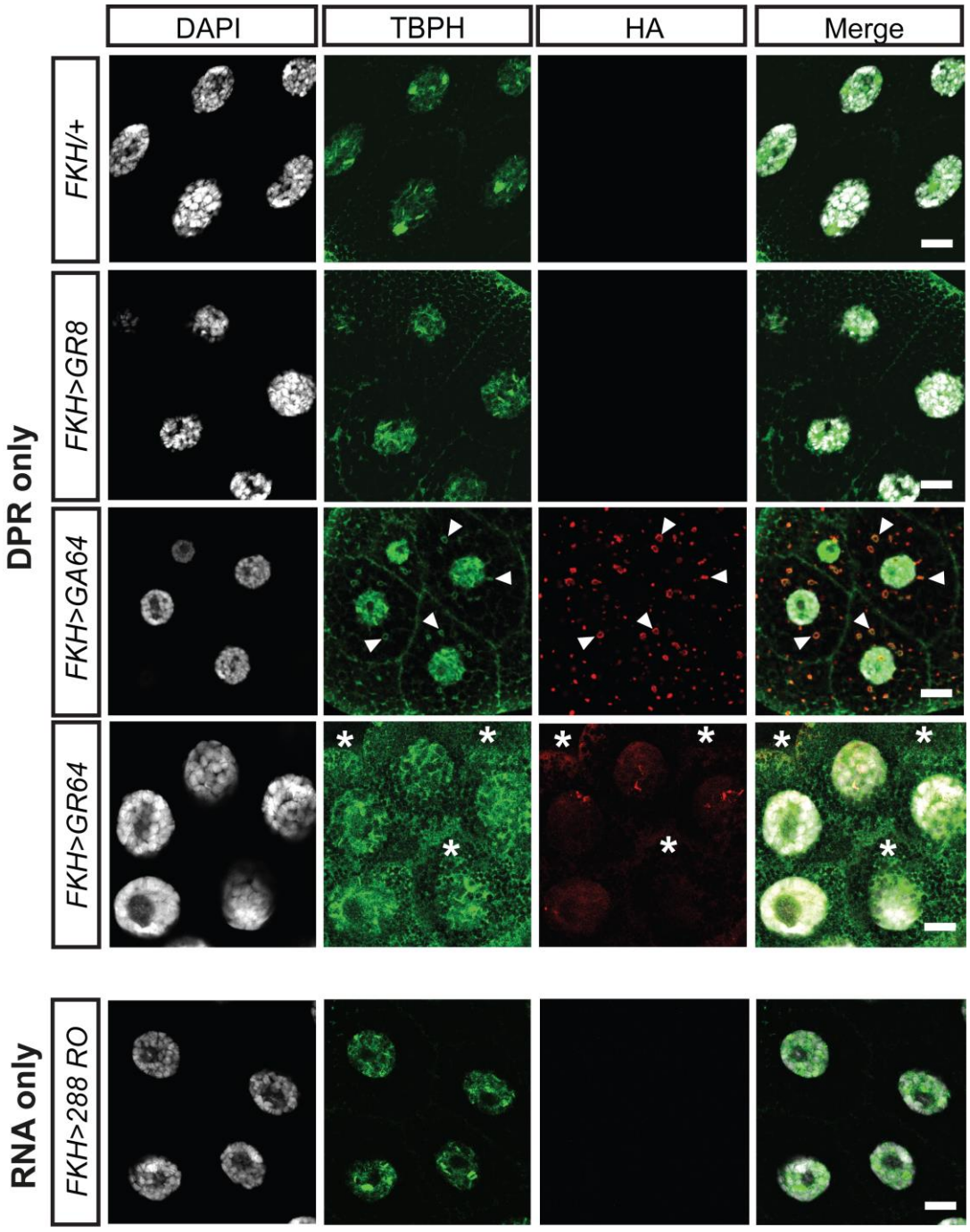

Figure 6.1 | DPR but not G_4C_2 RNA expression disrupts the localisation of *Drosophila* TDP-43. Salivary glands immunolabeled with *TBPH*, the *Drosophila* homologue of TDP-43. In controls (*FKH/+* and *FKH>GR8*) *TBPH* localisation is mainly nuclear. Compared to controls, poly-GA64 (*FKH>GA64*) expression causes cytoplasmic *TBPH* aggregates that co-localise with poly-GA inclusions (arrowheads). Poly-GR64 (*FKH>GR64*) expression leads to diffuse cytoplasmic mislocalisation of *TBPH* (asterisks) in all cases examined. Note, poly-GR64 expression also causes enlarged nuclei compared to controls. The expression of G_4C_2 RNA has no impact of *TPBH* localisation (FKH>288 RO). A minimum of 5 salivary glands were scanned per genotype; in each salivary gland 30 cells were investigated. Scale bars, 20µm.

6.2 Cytoplasmic TBPH enhances G4C2 toxicity

As nuclear depletion and cytoplasmic accumulation of TDP-43 is a characteristic feature of patients harbouring repeat expansions in *C9ORF72*, it was hypothesised that the cytoplasmic accumulation of *TBPH* may contribute to the early-onset phenotype in the 38 G_4C_2 repeat fly line that produces poly-GR. To investigate the impact cytoplasmic *TBPH* has on toxicity, a *TBPH* construct with a mutated NLS (*TBPH*ΔNLS) under control of the endogenous *TBPH* promoter was used (figure 6.2A) that causes cytoplasmic accumulation of *TBPH* in the cells flies harbouring the construct (appendix 7; figure A7). To determine whether elevating the levels of cytoplasmic *TBPH* within these flies would enhance the behavioural phenotype observed in the repeat flies, the *TBPH*ΔNLS construct was combined with 38 and 64 repeat flies and respective controls (figure 6.2B). The *TBPH*ΔNLS construct had no significant impact on climbing performance at day 5 when co-expressed with 64 G_4C_2 repeats that produces only poly-GP and poly-GA or when expressed in a control fly line (*Elav>ΔNLS/+*) (figure 6.2B) day 5. However, in the 38 G_4C_2 repeat line, (that in addition to poly-GP and poly-GA also produces poly-GR); *TBPH*ΔNLS significantly enhanced the climbing deficit seen at day 5 as measured by SING compared to the 38 repeats on their own (figure 6.2B). Hence co-expressing cytoplasmic TDP-43 is an enhancer of the G_4C_2 toxicity in flies.

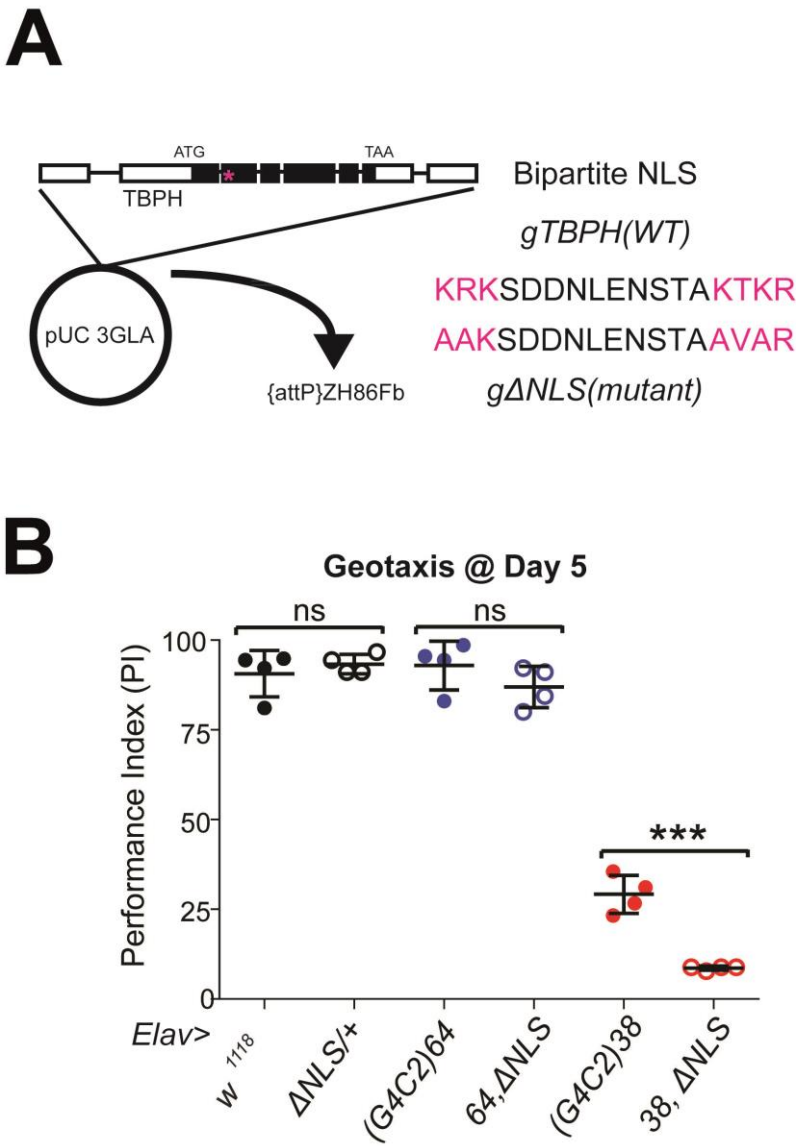

Figure 6.2 | Cytoplasmic TBPH enhances G4C2-related motor impairment. **(A)** Schematic of genomic *TBPH* region indicating exon (black boxes) intron structure, and protein coding start (ATG) and stop codon (TAA); pink asterisk indicates bipartite nuclear localisation signal (NLS) for wild type (WT) *TBPH* and mutant (ΔNLS) constructs subcloned into pUC 3GLA vector and inserted by site-specific integration at ZH86Fb on the third chromosome. **(B)** Day 5 motor performances of control flies and flies co-expressing ΔNLS-*TBPH* with 64 or 38 G4C2 repeats. A significant difference in the PI of genotypes tested at day 5 was observed using a one-way ANOVA ($F=218.6$, $p<0.0001$). Post-hoc analysis was performed using Bonferroni-Holm corrected multiple comparisons which can be found in appendix 10, table A10. Post-hoc analysis revealed co-expression of ΔNLS-TBPH had no impact on the day 5 motor performances of control flies or flies expressing 64 G4C2 at day 5. However, enhanced cytoplasmic TBPH causes a significant impairment of motor performance in flies expressing 38 G_4C_2 repeats compared to 38 repeat flies not expressing ΔNLS-TBPH ($p<0.0001$). Each data point represents the PI of ~15 flies. Mean with SEM is shown for each genotype (n=4).

6.3 Increasing cytoplasmic TBPH levels enhances translation of DPRs

TDP-43 has several functions in the cytoplasm (Ederle and Doorman, 2017) including regulatory roles in mRNA stability, mRNA trafficking and mRNA translation (Ederle and Doorman, 2017). These functions could be impacted by TDP-43 cytosolic accumulation. Of interest are the roles of TDP-43 in mRNA stability and mRNA translation. The expression of cytoplasmic TDP-43 has been shown previously to alter protein translation (Russo et al., 2017). Furthermore human TDP-43 has been shown to regulate the translation of peptides from a repeat expansion in a *Drosophila* SCA31 model (Ishiguro et al., 2017). Therefore it was speculated that enhanced cytoplasmic *TBPH* may have impact on DPR production in flies co-expressing both *TBPH*ΔNLS and G_4C_2 repeats. Flies containing G_4C_2 repeats that also co-expressed the *TBPH*ΔNLS were probed with poly-GP and poly-GA antibodies (figure 6.3). No DPRs were detected in any of the control fly lines (*Elav/+* and *Elav>ΔNLS/+*) (figure 6.3A and C). However, when both the 38 and 64 repeat constructs were combined with *TBPH*ΔNLS, the G_4C_2-*TBPH*ΔNLS flies demonstrated higher DPR expression levels compared to repeat containing flies not expressing *TBPH*ΔNLS for both poly-GP (figure 6.3A and B) and poly-GA (figure 6.3C and D). The effect was only not significant for poly-GP in flies expressing 38 repeats, however levels also appeared increased and significance may be reached with more replicates. Due to inabilities in accurately quantifying poly-GR levels with the current dot blot detection methodology this DPR was not probed for, although as demonstrated in chapter 3.8 of this thesis poly-GR was only detectable in the 38 repeat line at the timepoints probed for in this experiment. These western blots were performed in collaboration with Wing Hei Au.

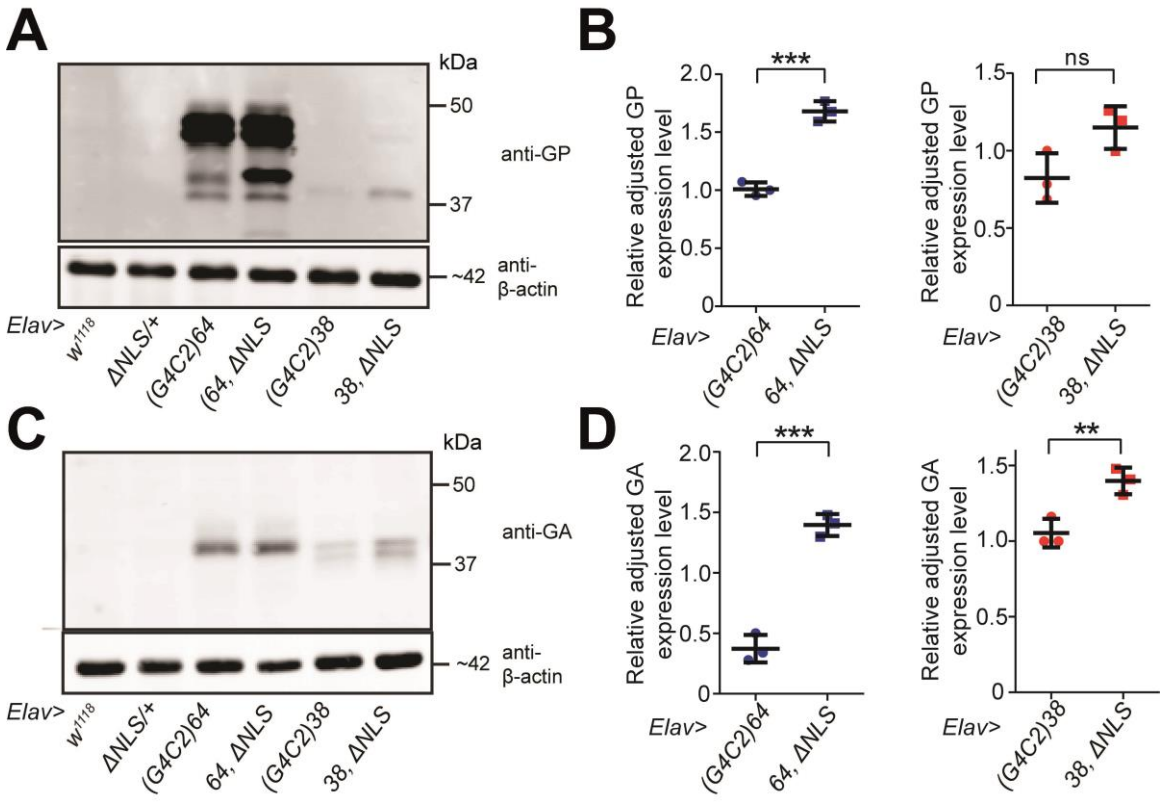

Figure 6.3 | Cytoplasmic TBPH enhances DPR levels. (**A**,**B**) Flies expressing 64 G4C2 repeats that also co-expressed the TBPHΔNLS construct demonstrated higher poly-GP levels compared to repeat containing flies not expressing TBPHΔNLS (t=11.02, ***;p<0.001). However, this effect was not significant in flies expressing 38 repeats. Mean and SEM are represented (n=3). (**C**,**D**) Flies expressing 64 and 38 G_4C_2 repeats also co-expressing the TBPHΔNLS construct demonstrated higher poly-GA levels compared to repeat containing flies not expressing TBPHΔNLS. (t=4.637, **; p<0.01), (t=12.15, ***;p<0.001). Mean and SEM are represented (n=3).

6.4 Discussion

6.4.1 DPR expression, but not G_4C_2 RNA, disrupts *Drosophila* TDP-43

The C9ALS/FTD post-mortem brain is characterized by the presence of RAN translated DPR inclusions and also cytoplasmic inclusions of the RNA binding protein TDP-43. Establishing a link between these two pathological hallmarks of the disease is a key question in the field that remains unanswered (Edbauer and Haass, 2016). The localisation of *Drosophila* TDP-43, *TBPH*, in control salivary gland cells (*FKH/+* and *FKH>GR8*) was predominantly nuclear with weak cytoplasmic *TBPH* also observed. *TBPH* staining was indistinguishable from controls in the G_4C_2-288 repeat RNA only line. Poly-GA64 expression did not alter this nuclear localisation of *TBPH*; however the protein formed deposits in the cytoplasm of salivary gland cells which co-localised which poly-GA64 inclusions. Poly-GR64 on the hand also did not cause a nuclear reduction in salivary gland cells; however its expression led to a striking increase in cytoplasmic *TBPH*.

Murine models have shown that expression of the G_4C_2 expansion is sufficient to cause TDP-43 pathology (Chew et al 2015, Liu et al 2016a). Chew et al. (2015) using a viral vector to express repeats in the mouse brain found phosphorylated endogenous mouse TDP-43 inclusions in the cortex and hippocampus of $(G_4C_2)_{66}$ mice but not $(G_4C_2)_2$ controls. Liu et al. (2016a) found *C9ORF72* BAC mice also harbour TDP-43 inclusions in degenerating neurons. Taken together these studies indicate the gain-of-function mechanisms derived from the expansion are sufficient to cause TDP-43 dysfunction. Human post-mortem work has revealed TDP-43 pathology appears after the formation of RNA foci and DPRs (Proudfoot et al., 2014; Baborie et al., 2015; Vatsavayai et al., 2016). Furthermore TDP-43 pathology is only present in end-stage BAC mice (Liu et al., 2016c). Thus the question is which of the toxic gain-of-function species produced from the G_4C_2 repeat expansion causes TDP-43 dysfunction (Edbauer and Hass, 2016). Some studies have suggested the G_4C_2 RNA may underlie TDP-43 mislocalisation. For example in the Chew et al. (2015) viral vector model cells with TDP-43 inclusions all contained at least one nuclear sense RNA foci. Cooper-Knock et al. (2015) found anti-sense RNA foci, but not sense RNA foci, correlate with TDP-43 mislocalisation in patient

motor neurons. In a G_4C_2 *Drosophila* model generated by Xu et al. (2013), reported to be RNA only, increased cytoplasmic *Drosophila* TDP-43 is observed (Zhang et al., 2015). However in the present study the expression of a 288 repeat RNA only G_4C_2 construct which forms foci in salivary gland cells does not cause any mislocalisation of *Drosophila* TDP-43. This suggests the G_4C_2 RNA is not the primary cause of TDP-43 dysfunction in C9ALS/FTD. This supports data from the DsRed2 repeat flies characterised earlier in which behavioural toxicity did not correlate with G_4C_2 RNA expression levels.

The salivary gland studies suggest rather that it is the DPRs that cause TDP-43 dysfunction. Poly-GR expression caused a large accumulation of *Drosophila* TDP-43 in the cytoplasm of salivary gland cells. Yamakawa et al. (2015) reported a co-localisation between TDP-43 and poly-GR/poly-PR inclusion bodies in the cytoplasm of HEK293T cells. Similar findings were reported by Lee et al. (2016) who observed a co-localisation between TDP-43, poly-GR/poly-PR and SG markers in the cytosol in HeLa cells. Mori et al. (2016) demonstrated cytoplasmic TDP-43 re-distribution when all three of the sense DPRs (poly-GP, poly-GA and poly-GR) were detected in a cell, but not with poly-GP or poly-GA alone or a combination of these DPRs (Mori et al 2016). The authors note in the replies to one of the reviewers of the manuscript that poly-GR accumulation likely contributes to TDP-43 mislocalisation; however they could not report on this as cells expressing poly-GR only were too rare quantify. Furthermore, poly-GA and poly-GR double-positive cells, but not cells with only poly-GA, frequently had cytosolic TDP-43 accumulation. Interestingly it was also observed that poly-GA inclusions often co-localised with *TBPH* in the cytoplasm of salivary gland cells expressing poly-GA64. Previous research has demonstrated a link between TDP-43 aggregation and poly-GA inclusions. For example in the human brain poly-GA inclusions have been shown to co-localise with TDP-43 aggregates; with poly-GA being in the centre of such inclusions suggesting it precedes TDP-43 aggregation (Mackenzie et al., 2013). Chew et al. (2015) found in their mouse model that ~75% of cells with TDP-43 inclusions also contained a poly-GA aggregate. Rare phosphorylated TDP-43 aggregates have also been reported in a poly-GA only mouse model. Khosravi et al. (2017) found cytosolic poly-GA inclusions disrupt the nuclear import of a reporter containing the NLS of TDP-43, inducing the partial mislocalisation of TDP-43 into cytoplasmic granules. Furthermore forcing poly-GA to

the nucleus rescued this. Hence the data from the present study further supports the link between poly-GA pathology and TDP-43 aggregation.

Hence these findings taken together strongly suggest that DPR accumulation, and not G_4C_2 RNA, causes TDP-43 dysfunction in C9ALS/FTD. Poly-GR expression caused a striking cytoplasmic accumulation of *Drosophila* TDP-43, whilst poly-GA expression led to the formation of cytoplasmic deposits of *Drosophila* TDP-43 which partially co-localised with poly-GA aggregates.

6.4.2 Cytoplasmic TBPH enhances G4C2 toxicity

In the 38 G_4C_2 repeat line that produces all 3 sense DPRs, expression of *TBPH*ΔNLS significantly enhanced the climbing deficit seen at day 5 as measured by SING compared to the 38 repeats expressed on their own. Hence cytoplasmic TDP-43 is an enhancer of the G_4C_2 toxicity in flies. Intestinally a cytoplasmic accumulation of *Drosophila* TDP-43 is seen following poly-GR expression which also shows early toxicity as evidenced by severe degeneration in the compound eye. Further, *Drosophila* TDP-43 inclusions that co-localise in the cytoplasm with poly-GA are seen as a result of poly-GA expression in salivary gland cells – but poly-GA does not cause any external degeneration when expressed in the *Drosophila* eye. This suggests a cytoplasmic redistribution of *TBPH* without the need for aggregate formation facilitates toxicity. Previous work has demonstrated that cytoplasmic TDP-43 is toxic and is itself sufficient to cause neurodegeneration. Barmada et al. (2010) found mutant TDP-43 toxicity correlated with a cytoplasmic accumulation of TDP-43 but not the formation of aggregates. Neither nuclear TDP-43 levels nor aggregates were predictors of neuronal cell death; cellular survival was strongly predicted by the amount of TDP-43 in the cytoplasm. Russo et al. (2017) found mutant TDP-43 with a non-functional nuclear localisation sequence (NLS) was toxic despite little aggregate formation. *In vivo* models have been developed to study the role of cytoplasmic TDP-43 in disease onset and progression. These models utilise either variants of TDP-43 that lead to cytosolic accumulation or mutant forms of the protein where the NLS has been mutated specifically so it is non-functional leading to the protein accumulating in the cytoplasm. Ritson et al. (2010) expressed forms of TDP-43 with either mutations in the NLS or nuclear export signal (NES) in the *Drosophila* eye in addition to wild-type human TDP-43 overexpression. Wild-type TDP-43 expression in the fly

eye produced a mild degenerative phenotype whilst the NES mutant TDP-43 produced no external eye phenotype. NLS-mutant TDP-43 expression however in produced a strong degenerative phenotype.

Evidence from mouse models also suggests TDP-43 aggregates are not required for toxicity and that elevated cytoplasmic levels of TDP-43 are sufficient in induce toxicity. Mice expressing human TDP-43 with a defective NLS (hTDP-43-ΔNLS) demonstrated progressive neurodegeneration in the neocortex, hippocampus and corticospinal tract. These mice also had a clasping motor phenotype - a measure of spastic motor impairment. Importantly, whilst cytoplasmic TDP-43 caused a loss of upper motor neurons leading to motor spasticity, only extremely rare cytoplasmic ubiquitinated pTDP-43 aggregates were observed; indicating cytoplasmic aggregation of TDP-43 is not required for its cytotoxicity and subsequent neurodegeneration (Igaz et al., 2011). hTDP-43-ΔNLS expression did however cause a reduction in endogenous mouse TDP-43 (mTDP-43) and neuronal cell loss correlated with the reduction of nuclear mTDP-43 staining. Arnold et al. (2013) found the toxicity of ALS-causing TDP-43 variants (TDP-43^{Q331K} and TDP-43^{M337V}) occurred without any nuclear loss of the human TDP-43 transgene or the formation of insoluble TDP-43 aggregates; however, TDP-43 was also not localised to the cytoplasm in these mice and neurodegeneration again correlated with the loss of endogenous mTDP-43 as in the Igaz et al. (2011) study.

Igaz et al. (2011) used a CaMK2α promoter which did not direct hTDP-43-ΔNLS expression to the spinal cord; hence the hTDP-43-ΔNLS mouse model did not develop an ALS-like phenotype. To overcome this the same group used a Dox-suppressible neurofilament heavy chain promoter for brain and spinal cord expression of hTDP-43-ΔNLS to further study the role of cytoplasmic TDP-43 in neurotoxicity (Walker et al., 2015). These mice developed progressive atrophy in cortical regions, neuromuscular junction and spinal cord in addition to severe muscular atrophy. The hTDP-43-ΔNLS mice also presented with several early onset ALS-like motor impairments and early death. Further studies with these mice revealed the expression of hTDP-43-ΔNLS also led to cognitive deficits hence mimicking FTD symptoms as well (Alfieri et al., 2014). Pathological phospho-TDP-43 (pTDP-43) accumulations were observed in 28% of neurons at end-stage but only 2 % spinal cord motor neurons had TDP-43 inclusions, despite this 50% of motor

neurons were lost in hTDP-43-ΔNLS mice by end stage further indicating cytoplasmic TDP-43 aggregates do not drive toxicity (Walker et al., 2015). Suppression of hTDP-43-ΔNLS also led to a phenotypic recovery which, although was accompanied by the removal of inclusions, cannot fully be attributed to this - as pTDP-43 inclusions were extremely rare in motor neurons as mentioned previously. Whilst the rescue of motor phenotypes by suppression of the hTDP-43-ΔNLS transgene supports cytoplasmic TDP-43 toxicity, these mice also had nuclear reduction of endogenous mTDP-43 which returned following removal of cytoplasmic human TDP-43 (Walker et al., 2015). Therefore, while this study and the Igaz et al. (2011) study point towards cytoplasmic TDP-43 aggregates not being major drivers of toxicity, both models are unable to distinguish whether the clearance of cytoplasmic TDP-43 and its suppression was responsible for the phenotypes and dramatic rescue seen or whether these were caused by the loss of nuclear TDP-43 and its subsequent return to the nucleus in the rescue experiment when hTDP-43-ΔNLS is suppressed.

Interestingly, even though mTDP-43 was cleared from the nucleus, Igaz et al. (2011) found a high proportion of the hTDP-43-ΔNLS protein was nuclear, suggesting a loss a nuclear TDP-43 per say is not entirely responsible for toxicity as the TDP-43-ΔNLS protein is also localised to the cytoplasm. This implies cytoplasmic TDP-43 has a direct neurotoxic effect itself (Igaz et al., 2011). Further characterisation of the CaMK2α hTDP-43-ΔNLS mice by Amile-Wolf et al. (2015) found that although endogenous mTDP-43 is lost; total nuclear TDP-43 protein levels are maintained in these mice. RNA seq analysis of the hTDP-43-ΔNLS mice revealed transcriptomic changes that were directly attributable to a gain of cytoplasmic TDP-43 function as comparisons with RNA-seq data from mice injected with a TDP-43 ASO revealed only little similarity. Furthermore, splicing changes were observed in the ASO mice but not the ΔNLS-hTDP-43 mice; indicating nuclear TDP-43 function is retained in these mice (Amile-Wolf et al., 2015). Hence the unchanged levels of nuclear TDP-43 combined with the large increase in cytoplasmic TDP-43 protein levels and minimal number of cytoplasmic TDP-43 aggregates strongly suggests a toxic gain of cytoplasmic TDP-43 function is responsible for toxicity and the changes in gene expression observed in the hTDP-43-ΔNLS mice.

Taken together the above results strongly suggest that TDP-43 is mislocalised to the cytoplasm is toxic and causes neurodegeneration. How might cytoplasmic TDP-43 cause toxicity? Microarray data from CaMK2α hTDP-43-ΔNLS mice revealed an increase in the levels of chromatin assembly and histone genes, suggesting a role for cytoplasmic TDP-43 in both transcription and histone transcript stability (Igaz et al., 2011). The RNA seq on these mice by Amile-Wolf et al. (2015) revealed cytoplasmic TDP-43 caused an upregulation of genes linked to transcription, assembly of chromatin, the nucleolus and protein metabolism. Cytoplasmic TDP-43 was also associated with a downregulation of genes with roles in synaptic activity and the posttranslational modification of proteins. Cytoplasmic TDP-43 caused a misregulation in the processing of histone 3'UTR genes. Deficits in both chromatin assembly and the nucleolus were seen in both the hTDP-43-ΔNLS mice and FTD-TDP post mortem tissue. The authors suggest a gain of cytoplasmic TDP-43 causes a dysregulation of nuclear dynamics. The cytosolic functions of TDP-43 in mRNA stability, mRNA trafficking and mRNA translation (Ederle and Dormann, 2017) could also all be impacted by TDP-43 cytosolic accumulation (Barmada et al., 2010).

In summary, clear evidence shows cytoplasmic TDP-43 mislocalisation is toxic independent of aggregate formation. The data presented shows interaction between the G_4C_2 expansion and a cytosolic mislocalisation of TDP-43; specifically, that it is an enhancer of G_4C_2 induced motor impairment. In poly-GR64 expressing salivary gland cells a strong cytoplasmic accumulation of TDP-43 is seen; it is tempting to speculate some of the poly-GR toxicity is mediated by this mislocalisation of TDP-43. One would predict if this cytoplasmic accumulation of TDP-43 does contribute to GR64 toxicity then reducing the cytoplasmic accumulation of TDP-43 would provide a rescue. Lee et al. (2016) found an RNAi to *TBPH* completely rescued the survival of poly-GR only flies to 100%, although this will also cause a nuclear reduction of the protein, it does suggest a toxic gain of TDP-43 function, rather than a loss per say, may play be involved in poly-GR toxicity.

6.4.3 Cytoplasmic *TBPH* enhances DPR levels

As discussed in the last section TDP-43 has several functions in the cytoplasm (Ederle and Doorman, 2017). Of interest are the roles of TDP-43 in mRNA stability and mRNA translation. If these functions are perturbed either the stability of the G_4C_2 mRNA could be affected or RAN translation of the G_4C_2 mRNA could be directly impacted – both possibilities could alter DPR levels. In the current study increasing cytoplasmic TBPH not only enhanced G_4C_2 mediated motor impairment, but also DPR levels. When the *TBPH*ΔNLS construct was combined with the both the 38 and 64 repeat constructs, and poly-GP and poly-GA levels examined, the amounts of both DPRs were significantly increased in flies co-expressing the *TBPH*ΔNLS compared to repeat flies without.

Several studies have now shown TDP-43 regulates the translation of certain mRNAs. Recombinant TDP-43 has been shown to behave as a translational repressor *in vitro* (Wang et al., 2008). TDP-43 also supresses the translation of *Rac1, GluA1*, and *Map1b* mRNAs via its interaction with the FMRP-CYFIP translational inhibition complex (Majumder et al., 2012; Majumder et al., 2016). In flies TDP-43 has been shown to regulate translation of *Futsch* mRNA at the neuromuscular junction (Coyne et al., 2014). Translational changes have also been reported in the motor neurons of a TDP-43 mouse model - with the RNA binding proteins DDX58 and MTHFSD being identified as misregulated (MacNair et al., 2016). Two separate TDP-43 proteomic studies discovered a large component of the TDP-43 interactome are proteins that regulate translation, including translation initiation factors and ribosomal proteins (Freibaum et al., 2010; Kim et al., 2010). In neurons cytoplasmic TDP-43 has been shown to form mobile RNA granules that contain mRNA binding proteins suggesting TDP-43 is involved in the transport and translation of these proteins (Alami et al., 2014). TDP-43 is also a component of stress granules which regulate translation during stress (Aulas and Vande Velde, 2015). Recently Russo et al. (2017) using a ΔNLS TDP-43 construct demonstrated cytoplasmic TDP-43 acts as global inhibitor of protein synthesis; the authors found TDP-43 RNA granules consist of polyribosomes which associate with the translational machinery. Overexpression of ΔNLS-TDP-43 led to increases in total amounts of the translation repressor 4E-BP1 and a decrease in the phosphorylation

of eIF4E; the consequence of which was repressed global protein translation. The interaction between cytoplasmic TDP-43 and the translational machinery was via the ribosomal protein RACK1. If cytoplasmic TDP-43 is a translational repressor, it is surprising that its co-expression with the G_4C_2 repeat constructs enhances DPR production in flies rather than inhibits it. Cytoplasmic TDP-43 suppression of translation has only been studied for canonical cap dependent ATG context (Russo et al., 2017); RAN translation of the G_4C_2 repeats involves non-canonical translation initiation mechanisms; whether cytoplasmic TDP-43 might affect canonical and non-canonical translation differently remains to be seen.

Another way in which enhancing the cytosolic levels of TDP-43 could impact the levels of DPRs is through changes in its interaction with the G_4C_2 RNA. Recently TDP-43 has been shown to play a vital role in regulating repeat associated translation in spinocerebellar ataxia type 31 (SCA31). SCA31 is neurological and neuromuscular disease caused by microsatellite pentanucleotide repeat expansions in an intron shared by the genes BEAN1 and TK2 (Sato et al., 2009). Ishiguro et al., (2017) focused on the TGGAA repeat expansion in SCA31 which is translated into poly-pentapeptide repeat (PPR) proteins. Both TDP-43 and FUS were found to bind to the UGGAA RNA repeats; further overexpression of human wt-TDP-43 suppressed the UGGAA rough eye phenotype in flies. TDP-43 interacted with UGGAA RNA via its RNA recognition motif and caused the UGGAA RNA foci produced to be smaller, although not degraded; suggesting TDP-43 acts as a chaperone for the folding of UGGAA RNA. However, remarkably, wt-TDP-43 co-expression reduced levels of the poly-WNGME PPR protein translated from the UGGAA expansion; further, this suppression was independent of the RNA recognition motif (RRM) of wt-TDP-43. Similar effects on RNA foci formation and PPR expression were seen for FUS and hnRNPA2B1.

The authors suggest that TDP-43 and other RNA binding proteins may act as chaperones for repeat RNA and propose three mechanisms as to how TDP-43 might reduce repeat associated translation. Firstly, by binding to the UGGAA expansion TDP-43 blocks the scanning ribosome, hence translation initiation. Secondly the RNA chaperone activity of TDP-43 changes the structural conformation of the UGGAA RNA required for translation of the repeats. Thirdly and finally the binding of TDP-43 to the UGGAA repeats in the nucleus prevents cytoplasmic transport and as

a consequence translation; in support of this MBNL1 expression represses repeat associated translation via a nuclear retention of repeat CUG/CAG RNA (Kino et al., 2015). Recent work by Zu et al. (2017) on RAN translation of the CCTG·CAGG expansion in DM2 supports the latter of these hypotheses; it was found that the translation of RAN proteins was modulated by the sequestration and nuclear retention of the repeat RNA by MBNL1. The TBPHΔNLS fly line used in this study not only causes a cytoplasmic accumulation of *TBPH* but also comes with a concomitant nuclear reduction of endogenous *TBPH* (Hirth lab unpublished data; appendix 7; figure A7). If this reduction of nuclear TBPH is sufficient to affect the normal nuclear function of the protein; then based on what was proposed by Ishiguro et al. (2017) for the interaction between TDP-43 and UGGAA mRNA; an increase in repeat associated proteins might be expected. For example TBPH would no longer be preventing the transport of G_4C_2 repeat RNA into the cytoplasm or no longer binding to the RNA in the nucleus in such a way that induces structural changes in the G_4C_2 repeat RNA unfavourable for translation. The data in this chapter presents evidence that endogenous fly *TBPH* can have the opposite effect on repeat associated translation reported by Ishiguro et al. (2017) and previous reports that cytoplasmic TDP-43 is a translation repressor (Russo et al., 2017). One explanation for this is that the interaction between TDP-43 and repeat RNA differs between different repeat sequences. TDP-43 preferentially binds to UG repeats hence its interaction with the UGGAA expansion in SCA31 is not unexpected (Polymenidou et al., 2011), further TDP-43 is found in the SCA31 RNA foci in human cerebellar purkinje neurons (Ishiguro et al., 2017). This suggests TDP-43 interacts differently with different repeat mRNAs. TDP-43 could for example induce a conformation change specific to the G_4C_2 sequence that promotes translation rather than decreases like for the UGGAA expansion. However, no direct binding has been shown between TDP-43 and G_4C_2 RNA in patient tissue and TDP-43 is not found to be sequestered into G_4C_2 RNA foci like other RNA binding proteins.

Intriguingly a partial knockdown of TBPH in the Ishiguro et al. (2017) UGGAA Drosophila model increased foci formation; but had no effect on repeat translated PPR protein levels. Furthermore when the RRM of TDP-43 was depleted PPR translation was still supressed. This observation suggests it is plausible that the increase in DPR expression seen as a consequence of TBPHΔNLS expression

could be attributable to some gain of toxic function resulting from enhanced cytosolic TDP-43, rather than a loss of nuclear function. The question is there any evidence for an interaction between cytosolic TDP-43 and the G_4C_2 RNA that could potentially promote translation from the G_4C_2 sequence? Ishiguro et al. (2016) found TDP-43 binds to G-quadruplex structures (G4), and is involved in the transport of G4 mRNAs to neurites where they are translated via the local protein translation complexes. This study suggests there is an important interaction between G-quadruplex mRNAs and cytosolic TDP-43 function in relation to intracellular trafficking of and local translation of G4 mRNAs. Could the enhanced DPR levels be due to increased translation at neurites that results from increasing cytosolic *TBPH* levels? DPRs are localised to axons (Hirth lab unpublished data; appendix 8; figure A8); further Burguete et al., (2015) found G_4C_2 RNA is localised in neurites where it induces branching defects and disrupts transport granule function. Hence it is not completely unfeasible to suggest DPRs could be translated locally in neurites.

It is also possible the increase in DPRs seen as a result *TBPH*ΔNLS expression is unrelated to the role of TDP-43 in regulating protein translation or a potential interaction with the G_4C_2 RNA. For example enhancing the levels of cytoplasmic TDP-43 may impact protein degradation or autophagy which could cause the DPRs to accumulate. Further DPR production has been shown to be increased under heat shock conditions in flies (Tran et al., 2015; Freibaum et al., 2015), hence increased cellular stress via an artificial increase in cytoplasmic *TBPH* could also cause an increase in DPR production. Irrespective of the causes, these results suggest a potential feedback loop in which the DPRs, specifically poly-GR, cause enhanced levels of cytoplasmic TDP-43, which itself further enhances the DPR levels leading to a cascade of events that ultimately culminates in neurodegeneration. Although levels of poly-GR were not probed for when the repeat lines were combined with the *TBPH*ΔNLS construct, due to difficulties in accurately quantifying and detecting this peptide, it is tempting to speculate that the enhanced phenotype seen for the 38 repeat line when co-expressing *TBPH*ΔNLS is due to increased neurotoxic poly-GR production.

6.4.4 Conclusion

The present study shows poly-GR on its own caused a cytoplasmic accumulation of *Drosophila* TDP-43, whereas poly-GA expression led to formation of inclusions of *Drosophila* TDP-43 in the cytoplasm. No changes in TDP-43 localisation or morphology were observed from the expression of a G_4C_2 RNA only construct. An amyloid cascade hypothesis for C9ALS/FTLD has been proposed where specific stressors cause TDP-43 to aggregate and become mislocalised to the cytoplasm (Edbauer and Haas, 2016). These results indicate the DPRs, and not the G_4C_2 RNA, is the key stressor in this cascade, accumulating early during disease progression and subsequently causing TDP-43 to be mislocalised to the cytoplasm. Furthermore cytoplasmic TDP-43 also enhances DPR production which excitingly suggests some kind of feedback mechanism where DPR mediated cytoplasmic TDP-43 accumulation enhances DPR production; potentially further exacerbating TDP-43 mislocalisation and dysfunction.

Chapter 7: Investigation of the localisation of nucleocytoplasmic transport proteins in flies expressing either poly-GR or poly-GA

Several studies have reported that both G_4C_2 RNA and DPRs impact nucleocytoplasmic transport (NCT), which may account for cellular toxicity observed in these C9 models (Zhang et al., 2015, Freibaum et al., 2015, Jovičić et al., 2015, Boeynaems et al., 2016a). Therefore I investigated whether the accumulation and aggregation of poly-GA and poly-GR affects NCT by looking at the localisation of important components of the NCT machinery which have also been identified as modifiers of G_4C_2 toxicity in previous studies (Zhang et al., 2015; Freibaum et al., 2015; Jovičić et al., 2015; Boeynaems et al., 2016a). Proteins involved in the shuttling of cargoes between the nucleus and cytoplasm and vice versa can roughly be divided into 3 categories (Cautain et al., 2015). Firstly, there are the nucleoporins (Nups) which build the nuclearpore complex (NPC); there are around 30 different vertebrate Nups in total that can be divided into 3 subgroups – the transmembrane Nups that anchor the NPC to the nuclear envelope, the FG-Nups which contain phenylalanine-repeats which selectively regulate nuclear transport, and finally the structural Nups that act as scaffold connecting the transmembrane Nups and the FG-Nups (Cautain et al., 2015). The second group of proteins are the nuclear transport receptors (NTRs). The NTRs constantly move between the cytoplasm and nucleus binding to their cytoplasmic or nuclear cargoes NLS and NES domains. The NTRs that bind to a proteins NLS domain and mediate their translocation into the nucleus are known as importins. NTRs that bind to proteins in the nucleus mediating their export from the nucleus into the cytosol are known as exportins (Cautain et al., 2015). The majority of NTRs are members of the karyopherin-β (KPNB) protein family (Cautain et al., 2015). Each member recognizes its own unique set of cargo proteins or RNAs (Cautain et al., 2015). KPNB proteins bind either directly or indirectly via adaptors to their protein and RNA cargoes. The adaptor proteins are specialized cargoes carrying an N-terminal NLS known as the importin β-binding domain (Pumroy and Cingolani, 2015). In the classic NLS pathway adaptor proteins bind to the NLS of the transport cargoes in the presence of KPNB1. The major class of adaptor proteins is the karyopherin-α (KPNA) protein family; there are 7 different

KPNAs divided into three subfamilies (Pumroy and Cingolani, 2015). The different subfamilies show selectivity for specific protein cargoes. Once formed the heterotrimeric KPNB/A/NLS-cargo complex shuttles through the NPC (Pumroy and Cingolani, 2015). The final component of the nuclear transport process is the Ran system. The shuttling of proteins into and out of the nucleus requires metabolic energy and this is provided by the small Ras-like GTPase Ran via GTP hydrolysis (Cautain et al., 2015).

TDP-43 is transported into the nucleus via the classic nuclear import pathway just described (Dormann and Haass, 2011). Defective nuclear import has been proposed to be the first step in the development of TDP-43 proteinopathy (Dormann and Haass, 2011), hence the morphology of proteins involved in the classic nuclear import pathway in the presence of poly-GR and poly-GA was investigated.

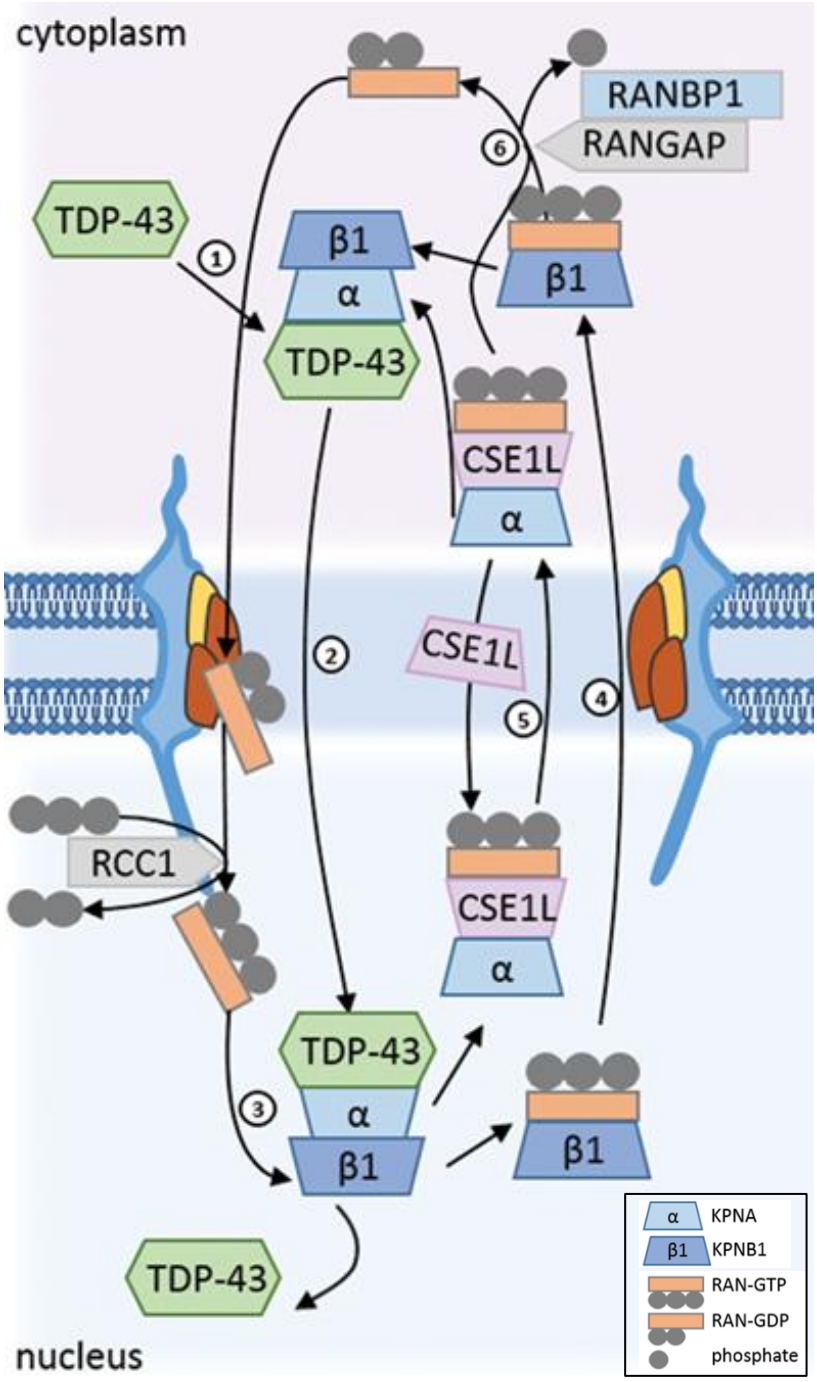

Figure 7.1 | Nuclear transport of TDP-43 via the classic nuclear import pathway. TDP-43 shuttles between the nucleus and cytoplasm via the classic nuclear import pathway. **(1)** The KPNAs (such as KPNA2 and KPNA4) bind to the nuclear localisation sequence (NLS) of TDP-43. **(2)** This KPNA-TDP-43 complex is then bound by KPNB1 in the cytoplasm forming the heterotrimeric KPNB/A/NLS-cargo complex shuttles through the NPC. **(3)** RCC1 is the nuclear exchange factor for RAN; converting RanGDP to RanGTP. RanGTP binding to KPNB1 causes the importin complex to dissociate thereby releasing TDP-43. **(4)** KPNB1 recycles back into the cytoplasm bound to RanGTP. **(5)** The KPNAs are recycled back into the cytoplasm together with CSE1L (also known as CAS) and RanGTP. **(6)** RanGAP together with RanBP1 increases the hydrolysis of RAN catalysing the conversion of RanGTP to RanGDP resulting in the dissociation of from KPNAs releasing them from the KPNA/CSE1L/RanGTP. Adapted from Mihevc et al., (2017).

7.1 Poly-GR or poly-GA expression does not cause a gross mislocalisation of proteins involved in the Ran system

Nucleocytoplasmic transport across the nuclear pore complex requires metabolic energy. The energy docking and undocking of transport cargoes is provided by the actions of small Ras-like GTPase RAN which exists in two nucleotide bound states; RanGTP and RanGDP. Energy for nuclear import comes from the GTP hydrolysis of RanGTP to RanGDP. RanGAP1 increases the hydrolysis of RAN inducing its GTPase activity; catalysing the conversion of RanGTP to RanGDP (Cautain et al., 2015). In its GDP bound state RAN dissociates from KPNAs releasing them from the KPNA/CAS/RanGTP complex allowing them to bind to their cytosolic cargoes (Stewart, 2007). Consistent with its function, RanGAP1 is localised on the cytoplasmic side of the nuclear pore (Fox and Tibbetts, 2015). *Drosophila* have one homologue of RanGAP1 known as *RanGAP*. The *FKH-GAL4* driver was used to express either poly-GR64 or poly-GA64 in salivary gland cells. This driver was chosen as the salivary gland cells have large nuclei and cytoplasm that can easily be distinguished. Salivary gland cells were stained with an anti-*RanGAP* antibody (Kusano et al., 2003) and anti-HA (figure 7.2A). In controls (*FKH/+* and *FKH>GR8*) RanGAP shows a smooth perinuclear localisation around the whole nuclear envelope as has been previously reported (Kusano et al., 2003) (figure 7.1A). This smooth perinuclear staining was also observed in 288-G_4C_2 RNA only (*FKH>288 RO*), GA64 (*FKH>GA64*) and GR64 (*FKH>GR64*) expressing salivary gland cells respectively (figure 7.2A). Hence no inclusions, cytoplasmic and nuclear puncta or sequestration of RanGAP into DPR aggregates was observed as has been reported in patient post-mortem tissue, viral vector $(G_4C_2)_{66}$ mice and a poly-GA only mouse model (Zhang et al., 2015; Zhang et al., 2016).

RCC1 (also known as RanGEF) is the nuclear exchange factor for RAN; converting RanGDP back to RanGTP. In the nucleus, RanGTP binds to KPNB, leading to conformational change in KPNB - causing the importin complex to dissociate thereby releasing the cargo (Stewart, 2007). Hence RCC1 has the opposite function to RanGAP1. RCC1 is bound to chromatin therefore is exclusively nuclear (Carazo-Salas et al., 1999). The *Drosophila* homologue of RCC1 is known as *Bj1*/*dRCC1*

(Frasch, 1991); *dRCC1* is required for the nuclear import of NLS cargoes (Shi and Skeath, 2004). *dRCC1* is diffusely nuclear in controls (*FKH/+* and *FKH>GR8*) (figure 7.2B) with extremely weak cytoplasmic staining. This nuclear staining was preserved in the poly-GA64 (*FKH>GA64*), poly-GR64 (*FKH>GR64*) and 288-G_4C_2 RNA only salivary glands (*FKH>288 RO*) (figure 7.1B) Some *dRCC1* was present in poly-GA64 inclusions consistent with previous reports in murine models demonstrating poly-GA sequesters proteins involved in nucleocytoplasmic transport into inclusions (Zhang et al., 2016). *dRCC1* staining was indistinguishable from controls in poly-GR64 expressing salivary gland cells – no decrease in *dRCC1* nuclear staining was observed as has been previously reported in *C9orf72*-ALS patient–derived neurons (Jovičić et al., 2015).

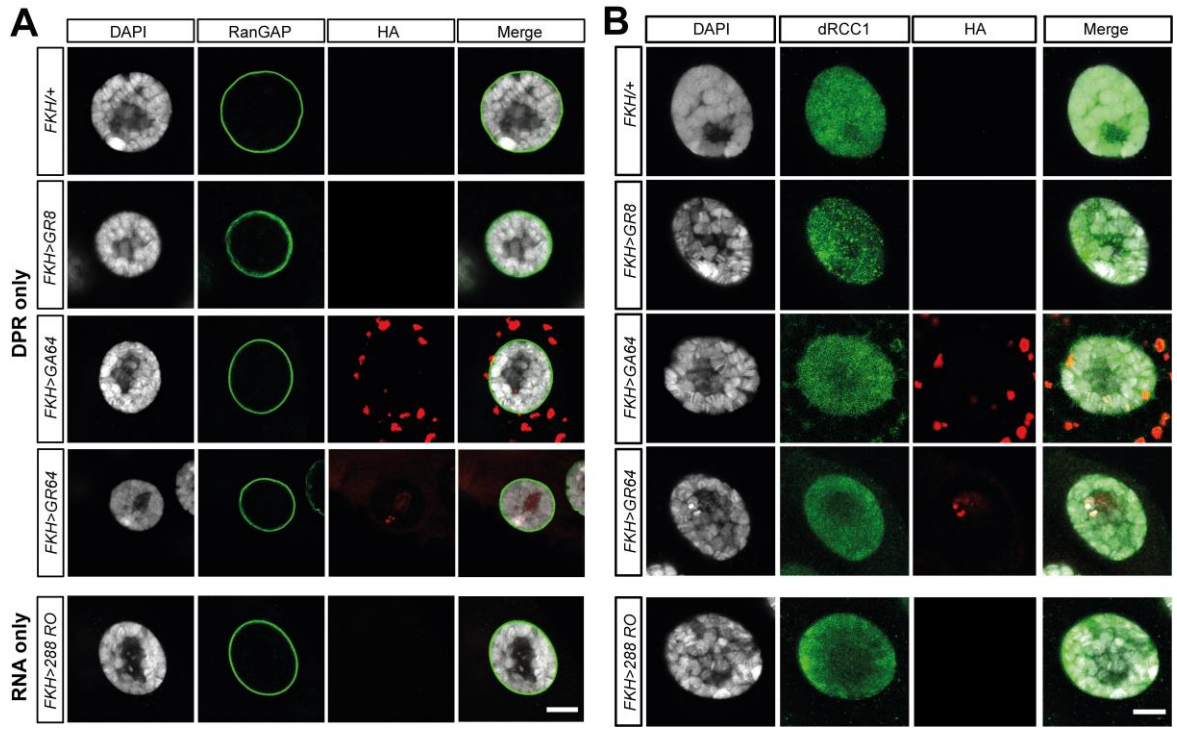

Figure 7.2 | Poly-GR and poly-GA accumulation does not alter the morphology of proteins involved in the Ran system. (A) Salivary glands were immunolabeled with *RanGAP*. No changes were seen in localisation of *RanGAP* in poly-GA64 (*FKH>GA64*), poly-GR64 (*FKH>GR64*) and RNA only (*FKH>288 RO*) expressing salivary glands compared to the controls (*FKH/+* and *FKH.GR8*). *RanGAP* was perinuclear in all genotypes. **(B)** Salivary glands were immunolabeled with *dRCC1*. No changes were seen in localisation of these proteins in poly-GA64 (*FKH>GA64*), poly-GR64 (*FKH>GR64*) and RNA only (*FKH>288 RO*) expressing salivary glands compared to the controls (*FKH/+* and *FKH>GR8*). *dRCC1* was nuclear in all genotypes; however co-localised with poly-GA inclusions in the cytoplasm. A minimum of 8 salivary glands were scanned per genotype. Scale bars, 10µm.

7.2 Poly-GR or poly-GA expression does not cause a gross mislocalisation of FG-repeat containing nucleoporins

The nucleoporins (Nups) build the NPC which regulates the flow of macromolecules from the cytoplasm to the nucleus and vice versa. Nups can be grouped into 3 categories based on structure and function; the transmembrane Nups, the structural Nups and the FG-Nups. The FG-repeat Nups were focused on as it has been recently shown that the anti-sense arginine DPR, poly-PR, binds to the central channel of the nuclear pore via an interaction with the polymeric form of the FG-repeat domain (Shi et al., 2017). Additionally, the FG-repeat Nup62 has recently been shown to be affected in both Huntington's disease mouse models and post mortem tissue (Grima et al., 2017). The following FG-repeat Nups were investigated MAB414 which labels several FG-Nups including Nup62, Nup153, Nup214, and Nup358 (figure 7.3A) and also *Drosophila Nup50* (figure 7.2B). The *FKH-GAL4* driver was used to express either poly-GR64 or poly-GA64 in salivary gland cells. MAB414 surrounded the nuclear envelope uniformly in the controls (*FKH*+ and *FKH>GR8*) (figure 7.3B). No gross morphological differences were observed between controls and the 288-G_4C_2 RNA only (*FKH>288 RO*) and poly-GA64 (*FKH>GA64*) salivary gland cells (figure 7.3B). Poly-GR64 (*FKH>GR64*) expressing salivary gland cells also showed a smooth distribution of MAB414 around the nuclear envelope (figure 7.3A). *Nup50* in controls was exclusively nuclear with an intense staining around the nuclear envelope (*FKH*>+ and *FKH>GR8*). This distribution was unaltered in 288-G_4C_2 RNA only (*FKH>288 RO*) salivary gland cells and also poly-GA64 (*FKH>GA64*) salivary gland cells despite cytoplasmic and nuclear aggregation of the latter (figure 7.3B). *Nup50* staining in poly-GR64 (*FKH>GR64*) salivary glands was also indistinguishable from controls despite nuclear poly-GR64 aggregation (figure 7.3B). In summary no obvious differences were seen in Nup morphology despite poly-GA, poly-GR or G_4C_2 RNA expression.

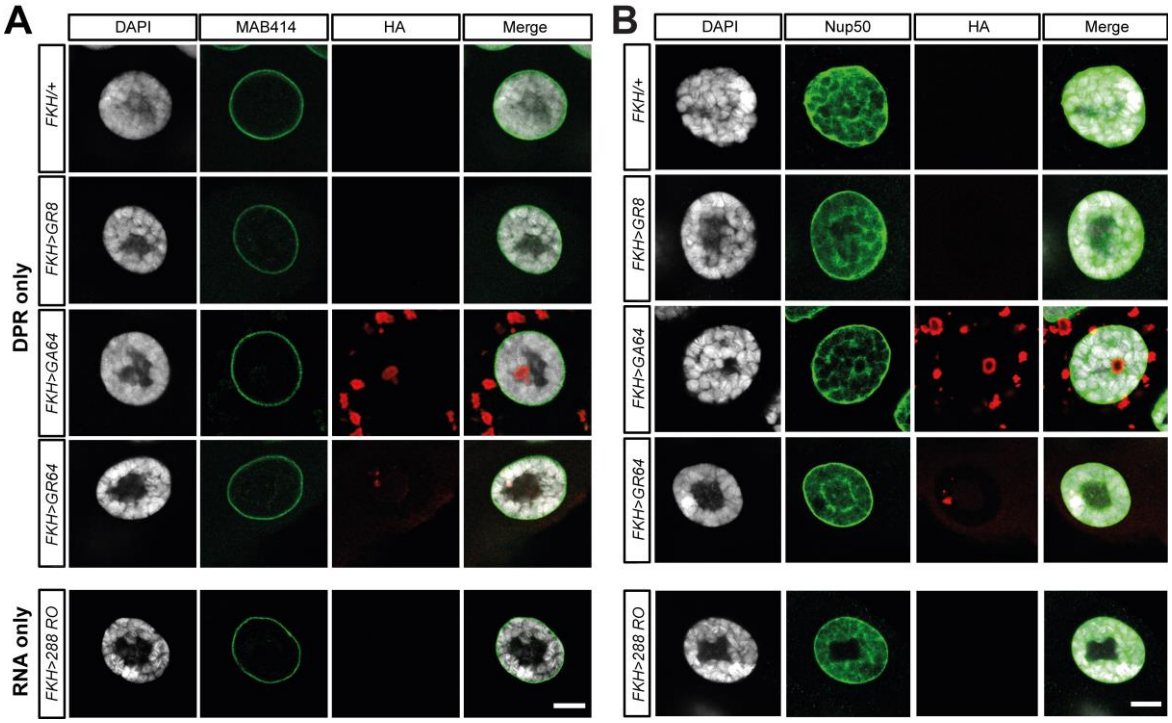

Figure 7.3 | Poly-GR and poly-GA accumulation does not alter the morphology of the nuclear pore complex. (A) Salivary glands were immunolabeled with MAB414 which labels FG-NUPs. No changes were seen in localisation of MAB414 in poly-GA64 (*FKH>GA64*), poly-GR64 (*FKH>GR64*) and RNA only (*FKH>288 RO*) expressing salivary glands compared to the controls (*FKH/+* and *FKH.GR8*). MAB414 was perinuclear in all genotypes. **(B)** Salivary glands were immunolabeled with *Nup50*. No changes were seen in localisation of these proteins in poly-GA64 (*FKH>GA64*), poly-GR64 (*FKH>GR64*) and RNA only (*FKH>288 RO*) expressing salivary glands compared to the controls (*FKH/+* and *FKH>GR8*). *Nup50* was nuclear in all genotypes with perinuclear enrichment. A minimum of 8 salivary glands were scanned per genotype. Scale bars, 10µm.

7.3 *Drosophila* karyopherin-α proteins are depleted from the nucleus and accumulate in the cytoplasm as a result of poly-GR expression

Proteins larger than 40 kDa in size are transported through the NPC via the NTRs (Cautain et al., 2015). The major NTRs are the KPNBs which bind either directly or indirectly to their protein cargoes (Cautain et al., 2015). TDP-43 is shuttled into the nucleus via the classic nuclear import pathway which involves the indirect binding of KPNB1 to TDP-43 via the KPNAs (Mihevc et al., 2017). Once in the nucleus this KPNB/KPNA/TDP-43 complex dissociates via binding of RAN-GTP to KPNB1; the RAN-GTP bound KPNB1 is then recycled back into the cytoplasm. KPNA cytoplasmic recycling is mediated via their binding with the nuclear export factor CAS also bound to RAN-GTP (Mihevc et al., 2017). KPNB1 and CAS are crucial to TDP-43 import; if they are knocked down KPNA recycling is diminished leading to the formation of cytosolic TDP-43 inclusions (Nishimura et al., 2010). The KPNA family of NTRs can be subdivided into three categories named α1, α2 and α3 (Pumroy and Cingolani, 2015). All 5 tested human KPNAs bind TDP-43; however KPNA3 and KPNA4, the 2 members of the α2 KPNA subfamily (Pumroy and Cingolani, 2015) are the strongest interactors with TDP-43 (Nishimura et al., 2010). *Drosophila* have homologues for the majority of human KPNA and KPNB NTRs (Chen et al., 2015). KPNA2 was identified as one of the strongest poly-GR and poly-PR interactors in a proteomics screen (Lee et al., 2016). Based on previous work suggesting KPNA dysfunction is key to TDP-43 aggregation (Nishimura et al., 2010), the localisation of the *Drosophila* homologue of KPNA4, *Importin*-α*3* and the *Drosophila* homologue of KPNA2, *Pendulin*/*Importin-α2* (a member of the α1 KPNA subfamily) was investigated (figure 7.4A and B). Unfortunately antibodies for the *Drosophila* homologues of KPNB1 (*Ketel*) (Lippai et al., 2000) and CAS (*Dcas*) (Tekotte et al., 2002) were either no longer avaliable or did not work, hence could not be studied in this thesis.

The *FKH-GAL4* driver was again used to express either poly-GR64 or poly-GA64 in salivary gland cells. *Importin*-α*3* in controls (*FKH/+* and *FKK>GR8*) is present in both the cytoplasm and nucleus although predominantly nuclear with enrichment around the nuclear envelope (figure 7.4A). This staining pattern is maintained in poly-GA64

(*FKH>GA64*) and 288-G_4C_2 RNA only salivary glands (*FKH>288 RO*). However, poly-GR64 (*FKH>GR64*) expression led to a nuclear depletion and cytosolic accumulation of *Importin-α3* (figure 7.4A). *Pendulin* in controls was present in both the nucleus and cytoplasm with an enrichment in the nucleus (*FKH/+* and *FKK>GR8*); again, this staining pattern was maintained in poly-GA64 (*FKH>GA64*) and 288-G4C2 RNA only (FKH>288 RO) (figure 7.4B). Poly-GR64 (*FKH>GR64*) expression on the other hand led to a dramatic nuclear depletion of *Pendulin* and cytosolic accumulation (figure 7.4B). Interestingly both proteins were present in some cytoplasmic poly-GA64 inclusions; furthermore, cytosolic levels of *Pendulin* appeared reduced in poly-GA64 flies which may be due to its sequestration into poly-GA inclusions.

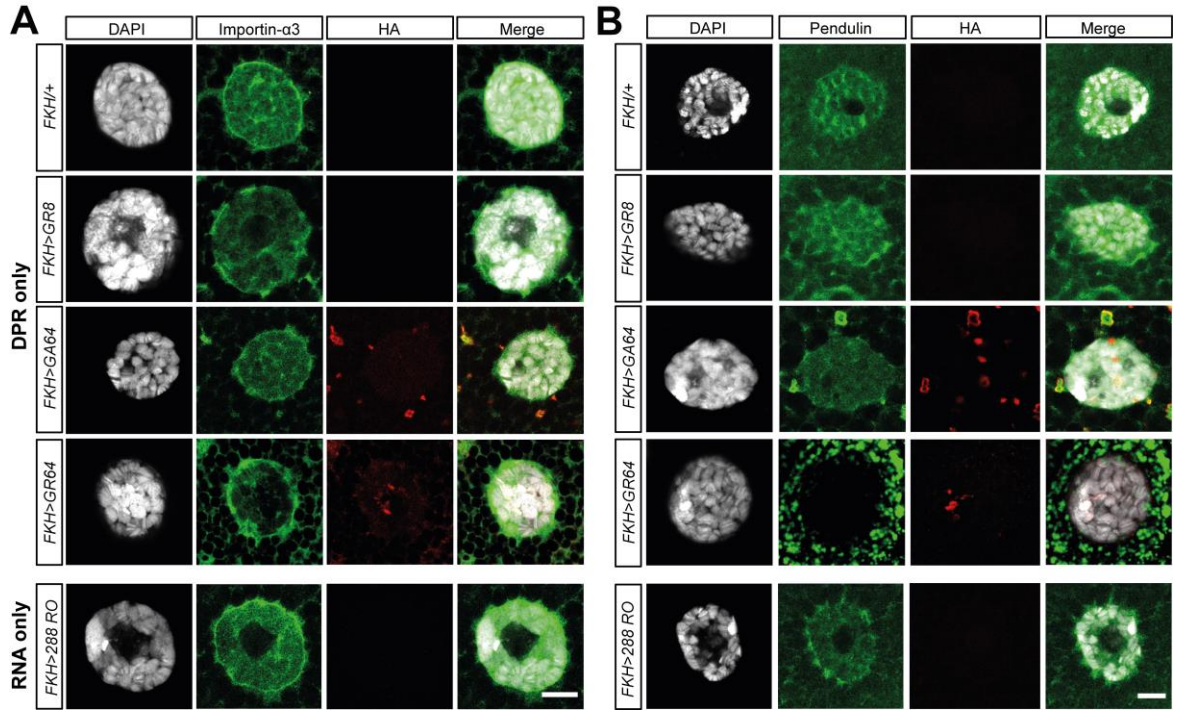

Figure 7.4 | Poly-GR causes a nuclear depletion of *Drosophila* KPNA2 and KPNA4. (A) Salivary glands were immunolabeled with *Importin-α3* (KPNA4). No changes were seen in localisation of *Importin-α3* in poly-GA64 (*FKH>GA64*) and RNA only (*FKH>288 RO*) expressing salivary glands compared to the controls (*FKH/+* and *FKH>GR8*). *Importin-α3* was predominantly nuclear with enrichment around the nuclear envelope in these genotypes, although *importin-α3* was sequestered into some poly-GA64 inclusions. However poly-GR64 (*FKH>GR64*) expression led to a nuclear depletion of *importin-α3*. **(B)** No changes were seen in localisation of *Pendulin* in RNA only (*FKH>288 RO*) expressing salivary glands compared to the controls (*FKH/+* and *FKH>GR8*). Poly-GA64 (*FKH>GA64*) expression however led to the sequestration of *Pendulin* into cytoplasmic poly-GA64 inclusions. Poly-GR64 (*FKH>GR64*) expression led to a nuclear depletion and cytosolic accumulation of *Pendulin*. A minimum of 8 salivary glands were scanned per genotype. Scale bars, 10µm.

7.4 Discussion

7.4.1 A nuclear depletion *Drosophila* KPNA2 and KPNA4, previously implicated in ALS/FTD, is observed as a consequence of poly-GR expression

Recent yeast, fly and mouse models have shown both nuclear import and export are disrupted by the presence of both the DPRs and G_4C_2 RNA (Zhang et al., 2015, Freibaum et al., 2015, Jovičić et al., 2015, Boeynaems et al., 2016a, Zhang et al., 2016). Important proteins involved in nucleocytoplasmic transport are clearly disrupted by the presence of G_4C_2 RNA and DPRs. In the present study poly-GR expression led to a nuclear depletion and cytoplasmic accumulation of the *Drosophila* homologues of KPNA2 and KPNA4. Poly-GA caused an aggregation of *Drosophila* KPNA2 and KPNA4 in the cytoplasm; these inclusions co-localised with poly-GA. No abnormalities in the NPC (as evidenced by staining for FG-repeat nups using MAB414 and Nup50 antibodies) or proteins involved in the Ran pathway (as evidenced by staining for the *Drosophila* homologues of RanGAP1 and RCC1) were observed as has been previously reported (Zhang et al., 2015; Freibaum et al., 2015). Some *dRCC1* was however present in poly-GA64 inclusions.

The RAN specific GTPase activating protein RanGAP1 had been shown to have increased nuclear localisation and form aggregates in *C9ORF72* repeat expansion post-mortem motor cortex in contrast to its normally smooth perinuclear staining in neurologically healthy controls (Zhang et al., 2015)**.** Increased nuclear localisation of the nucleoporin Nup205 was also observed in motor cortex cells by Zhang et al. (2015). Strong disruptions of the nuclear pore complex were also reported by Freibaum et al. (2015), the authors observed Nup107 in salivary gland cells had a distinct nuclear boundary but in cells expressing G_4C_2-58, the Nup107 staining was wrinkled and formed inclusions in close proximity to the nuclear envelope. Human iPSCs show increased cytoplasmic and reduced nuclear staining of RCC1, a Ran guanine nucleotide exchange factor (Jovičić et al., 2015). In a recent murine poly-GA only mouse model both RanGAP1 and Pom121, an inner membrane nuclear protein, were sequestered into poly-GA inclusions and were also mislocalised in mice expressing 66 G_4C_2 repeats from a viral vector (Zhang et al., 2016). RanGAP1, Nup88 and Nup62 all co-aggregate with the mutant Huntington protein in two

separate mouse models of Huntington disease (HD) (Grima et al., 2017). Furthermore RanGAP1 was found be aggregated in post-mortem HD striatum whilst Nup62 was mislocalised to either the cytoplasm or the nucleus in HD striatum (Grima et al., 2017). Additionally the HD-RAN proteins were also shown to impair nuclear import (Grima et al., 2017). Hence strong evidence implicates deficits in the NPC and Ran system as potentially causative of the nucleocytoplasmic transport defects.

However, in the DPR only poly-GR64 model used in this study, despite a cytoplasmic *TBPH* redistribution no gross mis-localisation or disruption of the NPC proteins, *RanGAP* or *dRCC1* was observed. For poly-GA no sequestration of *RanGAP* or NPC proteins into poly-GA aggregates was seen like reported by Zhang et al. (2016); although *dRCC1* was sequestered into cytosolic poly-GA inclusions. Both *RanGAP* and the NPC appear morphologically undisturbed, at least at the time points studied whilst *dRCC1* nuclear localisation was also preserved despite previous reports demonstrating the protein was depleted from the nucleus in C9 iPSCs by Jovičić et al. (2015). Not all studies have observed changes in the NPC or RanGAP; indeed a recent BAC mouse model that produced all 3 sense DPRs and demonstrated neuronal loss in the hippocampus in addition to cognitive deficits showed no mislocalised, discontinuous or punctate aggregates of RanGAP1 or Lamin B at 22 months of age (Jiang et al., 2016). Jovičić et al. (2015) saw no changes in RanGAP1 morphology in human C9ALS iPSCs. Although partial RanGAP1 mislocalisation was observed for poly-GA by Zhang et al. (2016), another poly-GA only mouse model reported no changes in RanGAP1. Similarly, Khosravi et al. (2017) found no evidence for altered Ran localisation or RanGAP1 co-aggregation with poly-GA inclusions in their cellular model; nor could the expression of either protein rescue the cytosolic accumulation of a TDP-43 reporter.

On the other hand, a striking nuclear reduction of the *Drosophila* homologues of KPNA2 (*Pendulin*) and KPNA4 (*Importin-α3*) was observed following poly-GR expression. Does this nuclear depletion and accumulation of KPNAs in the cytosol cause the cytoplasmic TDP-43 accumulation also observed as a consequence of poly-GR expression? Nishimura et al. (2010) found knockdown of individual KPNAs did not result in cytosolic TDP-43 accumulation; although the authors suggest that this may be because TDP-43 is imported into the nucleus by more than one KPNA. Indeed Nishimura et al. (2010) found 5 human KPNAs bind TDP-43 (KPNA1,

KPNA2, KNPA3, KPNA4, KPNA6 – although KPNA2 and KPNA4 were the strongest TDP-43 interactors). However knockdown of KPNB1 and CAS caused TDP-43 to accumulate in the cytoplasm and aggregate. Importantly knockdown of KPNB1 and CAS disrupted KPNA localisation; with CAS knockdown causing KPNA2 to accumulate in the nucleus and KPNB1 knockdown leading to a cytosolic accumulation and nuclear depletion of KPNA2 – like observed in this study for the *Drosophila* KPNA2 homologue *Pendulin*. This led Nishimura et al. (2010) to propose inhibition of nuclear import of KPNAs leads to cytosolic TDP-43 accumulation, as they directly bind TDP-43, but an initial disruption in KPNB proteins is key to KPNA dysfunction. Unfortunately *Drosophila* CAS and *Drosophila* KPNAB1 could not be tested. Poly-GR caused *Drosophila* TDP-43 to accumulate in the cytoplasm without an observable loss of nuclear *Drosophila* TDP-43. A nuclear depletion of TDP-43 may take time to develop and hence may not be observable in the early larval time point studied. Interestingly, analogous to what was reported for the effect of poly-GR on *Drosophila* TDP-43 localisation in this study, Nishimura et al. (2010) found KPNB1 knockdown caused a nuclear depletion and cytoplasmic accumulation of KPNA2, and, as just noted, this was accompanied by a concomitant cytoplasmic accumulation of TDP-43 in cells – however, despite this cytosolic increase in TDP-43, no loss of nuclear TDP-43 was seen in KPNB1 siRNA treated cells.

How might poly-GR disrupt KPNA localisation? Poly-GR is known to cause oxidative stress (Lopez-Gonzalez et al., 2016) and previous work has demonstrated the KPNAs are also mislocalised in cells exposed to oxidative stress (Kodiha et al*.,* 2008). Poly-GR also causes poorly dynamic SGs that do not disassemble in the cytoplasm, the KPNAs have been shown to be SG components. KPNA2 is found abundantly in SGs following arsenite treatment (Fujimura et al., 2010). Mahboubi et al. (2013) expanded this further finding all three of the KPNA subfamilies (α1, α2 and α3) are components of SGs under stress conditions. This recruitment appeared specific to SGs as the KPNAs were not found in P-bodies. Mahboubi et al. (2013) and Chang and Tarn. (2009) found KPNB family members, which bind the KPNAs in the cytoplasm are recruited to SGs however this was not the case for CAS which binds the KPNAs in the nucleus (Mahboubi et al., 2013). The KPNAs and KPNB were present in more than 90% of SGs in total. The presence of KPNA and KPNB proteins in SGs has led to the suggestion that they are also involved in SG assembly

(Mahboubi et al., 2013). Mahboubi et al. (2013) suggest KPNA and KPNB proteins may be involved in transporting RNAs and proteins to the growing SG. Hence it will be important to look if KPNA and KPNB proteins are present in the poorly dynamic SGs induced by poly-GR (Lee et al., 2016). Deficits in the NPC resulting from poly-GR expression may also be involved in the cytoplasmic accumulation of the KPNAs observed. Just because no morphological deficits were observed this does not mean the function of the NPC is not impaired; indeed poly-PR, binds to the FG-repeats in the central channel of the pore making the NPC less permeable. Poly-GR may act through similar mechanisms meaning the KPNAs are unable to diffuse through the NPC. Indeed, as just noted indirect support that a combination of NPC and KPNA dysfunction could be important for TDP-43 mislocalisation comes from Khosravi et al. (2017) who found overexpression of NPC proteins, KPNA3 and KPNA4 but not Ran or RanGAP1 rescues cytosolic TDP-43 accumulation. Poly-GR also strongly interacts with KPNA2/*Pendulin* hence some aspect of this direct interaction may also disrupt the localisation of this nuclear import adaptor protein (Lee et al., 2016).

Both KPNA2 and KPNA4 *Drosophila* homologues were sequestered into poly-GA inclusions in the cytoplasm consistent with previous reports in murine models demonstrating poly-GA sequesters proteins involved in nucleocytoplasmic transport into inclusions (Zhang et al., 2016). A link between TDP-43 dysfunction, poly-GA and the KPNAs has recently been shown by Khosravi et al. (2017). The authors observed poly-GA expression causes a TDP-43-NLS reporter to be recruited cytoplasmic granules in transfected hippocampal neurons. Khosravi et al. (2017) found KPNA/B-mediated nuclear import, the import pathway involved in TDP-43 shuttling, is perturbed by poly-GA expression. Poly-GA inhibited the nuclear import of the KPNA4 cargo p65. In addition poly-GA specifically inhibited KPNA/B-mediated nuclear import as poly-GA had no impact on TNPO mediated nuclear import which involves the binding of TNPO to PY-NLS motifs. The localisation of a reporter containing the PY-NLS of hnRNPA1 was not affected by poly-GA expression. Finally Khosravi et al. (2017) observed overexpression of proteins involved in importin α/β-mediated shuttling of TDP-43 such as KPNA3, KPNA4 and NUP62, but not Ran or RanGAP1, restored nuclear localisation of the TDP-43-NLS. How might poly-GA cause cytoplasmic aggregation of KPNAs? Poly-GA amyloidogenic fibrils form a parallel β-sheet structure; Woerner et al. (2016) demonstrated in a landmark paper

that cytoplasmic β proteins sequester nucleocytoplasmic transport proteins – including KPNA2 and KPNA4 – into cytoplasmic inclusions. This was specific to cytoplasmic β proteins aggregates, as nuclear β proteins aggregates did not impact the localisation of nucleocytoplasmic transport proteins. In support of this Khosravi et al. (2017) found rerouting poly-GA aggregates into the nucleus prevented cytoplasmic mislocalisation of TDP-43. Hence the secondary structure of poly-GA may underlie the sequestration of KPNA2 and KPNA4 into poly-GA inclusions. TDP-43 directly binds to KPNA2 and KPNA4 (Nishimura et al., 2010) hence the presence *Drosophila* TDP-43 in poly-GA inclusions could be due to its direct binding to these nuclear adaptors in the cytoplasm.

Finally no changes in the localisation of nucleocytoplasmic transport proteins was observed following the expression of a G_4C_2 RNA only construct which does not produce DPRs. This argues the disrupted nucleocytoplasmic transport observed in G_4C_2 fly models is attributable to the actions of the DPRs, rather than the G_4C_2 RNA (Zhang et al., 2015; Freibaum et al., 2015).

7.4.2 Conclusion

In summary an investigation of the different components involved in nucleocytoplasmic transport, including proteins involved in the Ran pathway (RanGAP1 and RCC1) and the nuclear pore complex (FG-repeat containing Nups) revealed the *Drosophila* homologues of nuclear import adaptor proteins, KPNA2 (*Pendulin*) and KPNA4 (*Importin-α3*), involved in the nuclear import of TDP-43, are depleted from the nucleus and accumulate in the cytoplasm following poly-GR expression, suggesting a loss of KPNA function. Interestingly although *Drosophila* KPNA2 and KPNA4 are both sequestered by poly-GA into aggregates they are not depleted from the nucleus like what is observed following poly-GR expression suggesting a failure in nuclear import of the KPNAs is important to the phenotype observed in the poly-GR salivary gland cells. An inhibition of the nuclear import of KPNAs has been proposed previously to lead to cytosolic TDP-43 accumulation (Nishimura et al., 2010). Hence a loss of KPNA function may be responsible for the cytoplasmic *TBPH* accumulation observed following poly-GR expression.

Chapter 8: Human KPNA4 is affected in both sporadic FTD-TDP and C9FTD/ALS post-mortem tissue and is sequestered into pathological TDP-43 inclusions

TDP-43 is transported into the nucleus via this classic nuclear import pathway (Nishimura et al., 2010). The localisation *Importin-α3*, the *Drosophila* homologue of KPNA4 and *Pendulin*, the *Drosophila* homologue of KPNA2, are disrupted via poly-GR expression. KPNA4 and KPNA2 directly bind TDP-43 and are involved in its nuclear transport (Nishimura et al., 2010). KPNA2 nuclear staining was previously shown by Nishimura et al. (2010) to be reduced in FTD-TDP frontal cortex compared to controls; furthermore, cytoplasmic KPNA2 staining was also absent in FTD-TDP neurons. KPNA2 levels were also downregulated in FTD-TDP frontal cortex tissue. KPNA2 cytoplasmic staining was increased in ALS spinal cords compared to controls. To investigate whether KPNA4 expression and localisation is also disrupted in TDP-43 proteinopathies, the abundance and distribution of KPNA4 and KPNA2 was investigated in the frontal cortex of C9FTD and sporadic TDP-43 positive FTD (FTD-TDP) patients. Details on the cases used for the human post-mortem study can be found in appendix 9; table A9.1.

8.1 Soluble protein levels of KPNA4 are downregulated in SFTD-TDP and C9FTD frontal cortex and present in the detergent-insoluble UREA fraction

To determine whether nuclear import may be perturbed in sporadic FTD-TDP and C9FTD, western blotting was performed on frontal cortex lysate from controls, sporadic FTD-TDP and C9-FTD cases (figure 8.1A & B). Membranes were probed for the *Drosophila* homologue of *Importin-α3* – KPNA4. The protein levels of the RIPA soluble fraction of KPNA4 were lower in both sporadic FTD-TDP and C9FTD frontal cortex samples compared to neurologically healthy controls (figure 8.1A). Although there was a clear trend showing soluble levels of KPNA4 were lower in the disease brain - this was not statistically significant likely due to the low number of samples probed (n=4-5) (figure 8.1B). Two of the disease samples (lanes 8 and 11, figure 8.1B) show no KPNA4 signal. A vital element of postmortem research is tissue quality, which unlike animal tissue – can not be controlled accurately at death (Stan et al., 2006). Numerous pre- and postmortem factors influence tissue quality which in turn influences the accuracy of results (Stan et al., 2006). To determine whether the lack of KPNA4 in these samples is an accurate representation of KPNA4 levels or a confound due to poor tissue quality, the standard of these two cases should be assessed using measures such as tissue pH, markers of RNA quality (such as RNA Integrity Number) and several measures of protein quality.

Despite this the decrease in the soluble levels of KPNA4 suggests the protein could be accumulating in the insoluble fraction, indicating aggregate formation. Aggregation prone disease related proteins such as tau and amyloid-β are insoluble in strong-detergent buffers but can be solubilised by urea buffer and analysed on SDS-PAGE (David et al., 2010). To explore the possibility that KPNA4 might have altered solubility in disease tissue, RIPA extracted insoluble pellets from homogenised frontal cortex samples were treated with urea buffer to isolate the detergent insoluble proteins from the pellet. No KPNA4 signal was seen from the urea soluble fraction from control cases. However, for sporadic FTD-TDP and C9-TDP, KPNA4 was present in the urea fraction, with a high molecular weight band seen in the stacking gel; especially in C9FTD cases (figure 8.1A). This suggests the protein could be forming insoluble inclusions in patient brains.

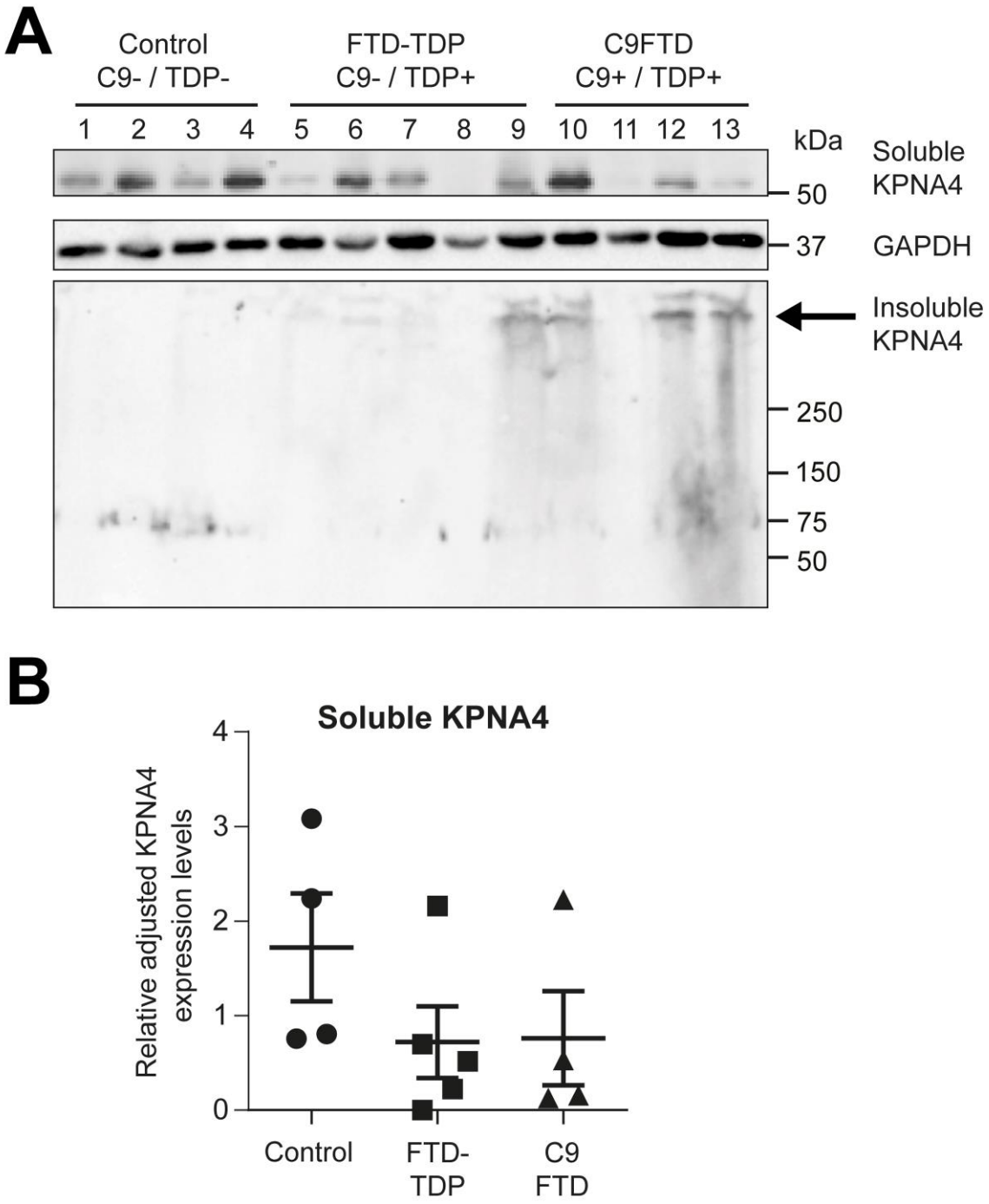

Figure 8.1 | Western blot analysis of extracts from post-mortem frontal cortex of patients with sporadic FTD-TDP, C9FTD and control brain samples. (A & B) Accumulated KPNA4 is found in the insoluble fraction of sporadic FTD-TDP and C9FTD samples (black arrow), but not in controls. **(B)** Quantitative western blotting soluble KPNA4 indicates down-regulation in both conditions however this was not statistically significant (F=1.377, p=0.2964) with the current N (4-5 cases per condition).

8.2 Immunolabeling reveals mislocalisation and aggregation of KPNA4 in sporadic FTD-TDP and C9FTD frontal cortex

The presence of KPNA4 in the urea fraction suggests the protein may be forming inclusions in the sporadic FTD-TDP and C9FTD brain. To investigate localisation of KPNA4 in disease frontal cortex sections from control, sporadic FTD-TDP and C9-FTD cases were probed with a KPNA4 specific antibody. KNPA4 immunolabeling in controls showed a uniform distribution in neurons with both nuclear and cytoplasmic distribution (figure 8.2 a, b, c, d). In sporadic FTD-TDP cases, there are two different staining patterns; either increased nuclear staining accompanied by a lesser intensity of cytoplasmic labelling (figure 8.2 e, f) or nuclear clearing with preserved or mildly increased cytoplasmic positivity (figure 8.2 e, g). In C9FTD cases there was increased KPNA4 immunoreactivity compared to controls and sporadic FTD-TDP (figure 8.2A i). In addition KPNA4 staining in C9FTD revealed cells with absent nuclear staining and harbouring immunoreactive cytoplasmic inclusions (figure 8.2 k). Further positivity in neuronal cell processes consistent with axons and dystrophic neurites was also observed in C9FTD frontal cortex. (figure 8.2l). Interestingly, in both sporadic FTD-TDP and C9FTD brains nuclear depletion was often accompanied by nucleolar immunoreactivity (figure 8.2 f, j). Semi-quantitative analysis of KPNA4 DAB immunolabeling was performed by neuropathologist Dr. Tibor Hortobagyi (KCL). KPNA4 localisation was further probed via immunofluorescence. KPNA4 was distributed between the cytoplasm and nucleus in controld (figure 8.3 a, b, c; figure 8.4) as has been previously reported in human spinal cord (Latiman et al., 2017) and cell lines (Woerner et al., 2016). In the sporadic FTD-TDP (figure 8.3 d, e, f; figure 8.4) and C9FTD (figure 8.3 g, h, I; figure 8.4) cases KPNA4 was completely depleted in a number of some neurons. Hence these data suggest KPNA4 is mislocalised in FTD TDP-43 proteinopathies. A semi-quantitive analysis of the frequency with which this KPNA4 mislocalisation was observed is present in appendix 9 table A9.2.

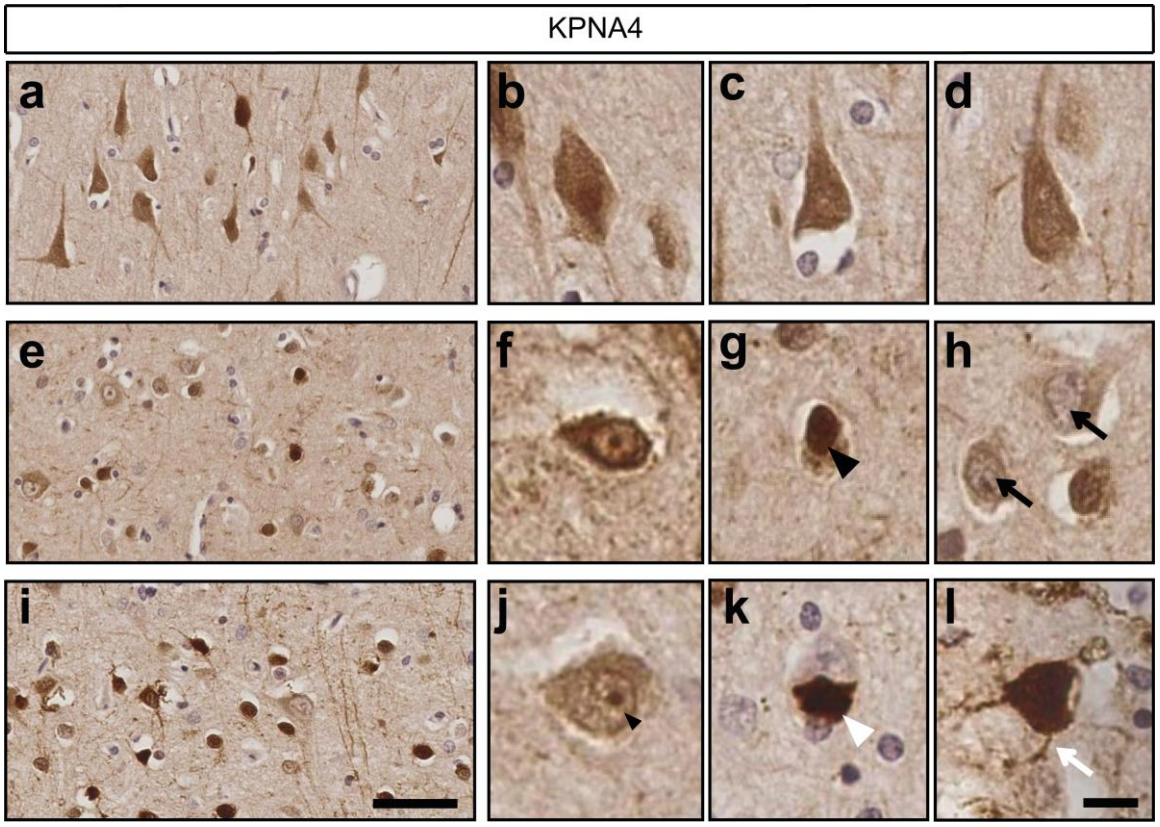

Figure 8.2 | KPNA4 localisation is disrupted in the post-mortem frontal cortex of patients with sporadic FTD-TDP, C9FTD but not control brain samples. (a-d) DAB immunolabeling in controls showed a uniform KPNA4 distribution in neurons with both nuclear and cytoplasmic localisation. **(e-h)** In sporadic FTD-TDP cases, cells lacking nuclear KPNA4 immunolabeling show nucleolar immunoreactivity **(e**, **f)**. Cells with increased nuclear staining show decreased cytoplasmic labelling **(e**, **g**, black arrowhead**)**, whereas cells with nuclear depletion show preserved or mildly increased cytoplasmic immunolabeling **(e**, **h**, black arrows**)**. **(i-l)** In C9ALS/FTD cases, KPNA4 DAB immunolabeling in the nucleolus is also observed compared to sporadic FTD-TDP and control cases **(j)**. Numerous cells with absent nuclear KPNA4 immunostaining display immunoreactive cytoplasmic inclusions **(k**, white arrowhead**)** and also **(l)** immunolabeled neuronal cell processes indicative of dystrophic neurites (white arrow). Scale bars, 50μm **(a,e,i)**, 10μm **(b,c,d,f,g,h,j,k,l)**.

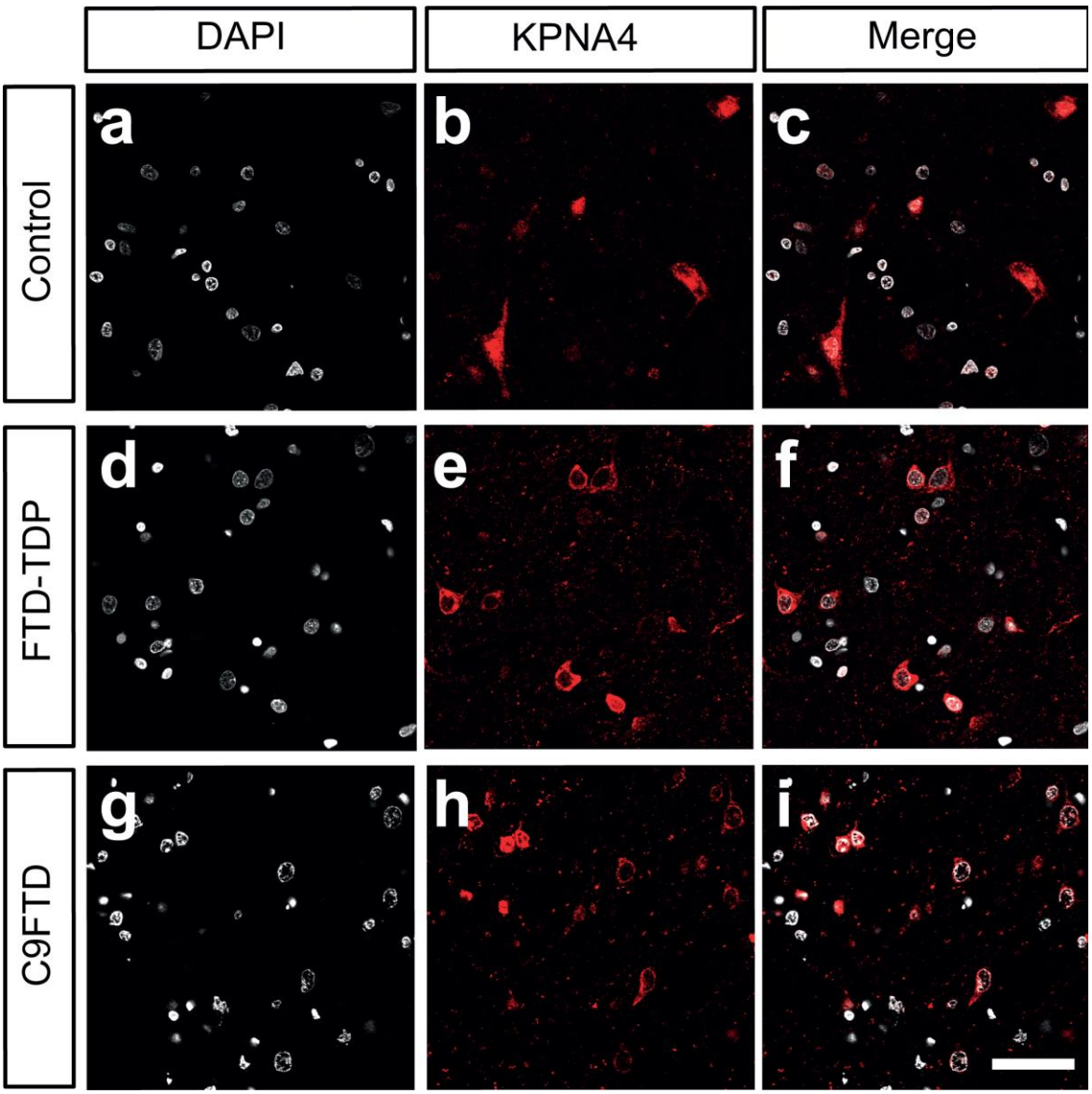

Figure 8.3 | Nuclear depletion and cytoplasmic accumulation of KPNA4 in sporadic FTD-TDP and C9FTD human frontal cortex. (a-c) KPNA4 immunolabeling in neurons of control cases reveals uniform distribution in both nucleus and cytoplasm. **(d-f)** In sporadic FTD-TDP cases, immunolabeling reveals nuclear depletion and cytoplasmic accumulation of KPNA4. **(g-i)** In C9FTD cases, immunolabeling reveals further pronounced nuclear depletion and cytoplasmic accumulation of KPNA4. Scale bar, 50μm **(a-i)**.

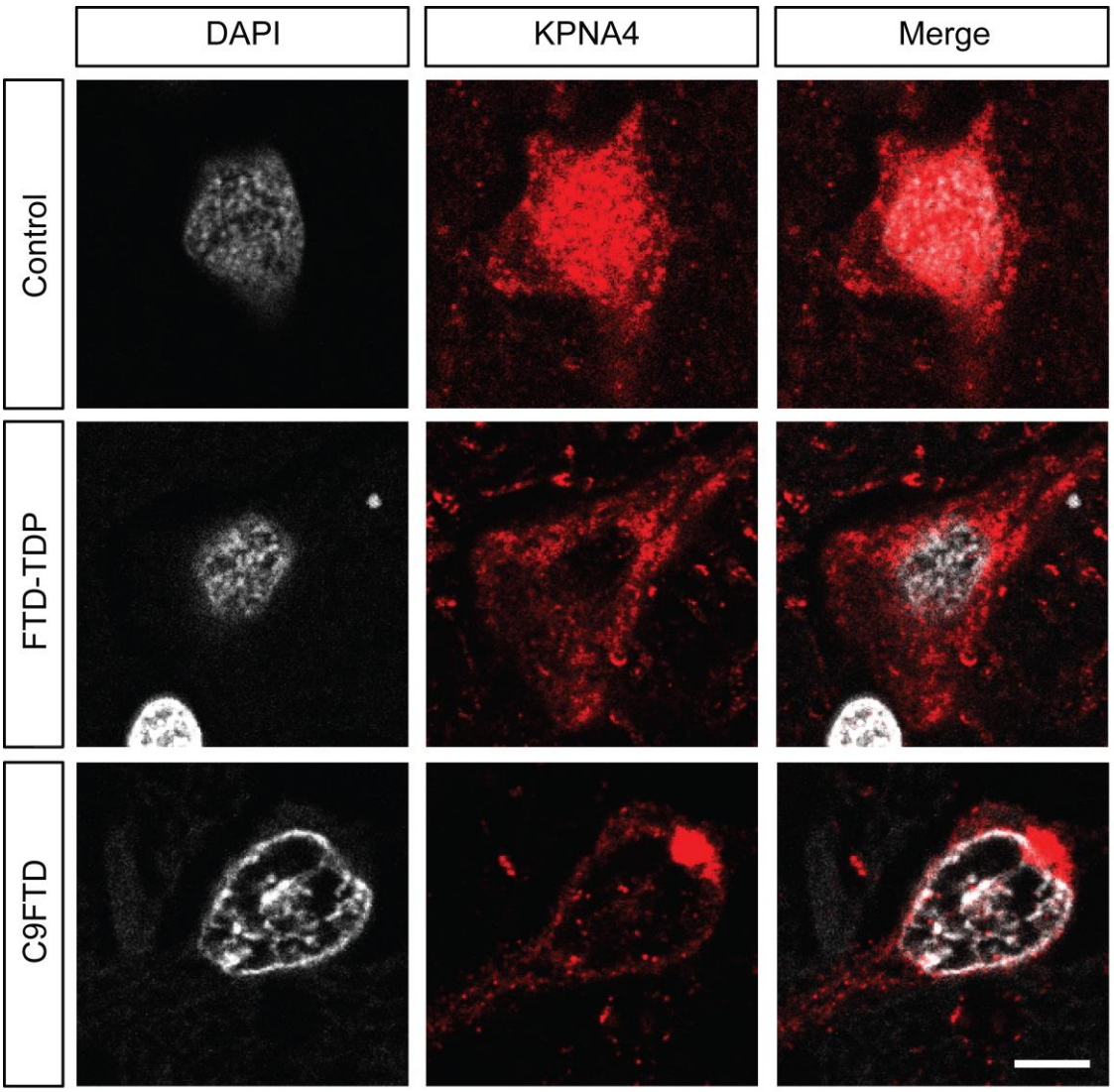

Figure 8.4 | Nuclear depletion of KPNA4 in sporadic FTD-TDP and C9FTD human frontal cortex. KPNA4 immunolabeling in neurons of control cases reveals a uniform distribution in both nucleus and cytoplasm. In sporadic FTD-TDP cases, a nuclear depletion and cytoplasmic accumulation of KPNA4 was observed. In C9ALS/FTD cases, a further pronounced nuclear depletion and cytoplasmic accumulation of KPNA4 was seen. Scale bar, 10µm.

8.3 KPNA4 overlaps with phospho-TDP-43 inclusions in patient brains

Insoluble aggregates of the RNA binding protein TDP-43 is the pathological hallmark of C9ALS/FTD (Mackenzie et al., 2014). Using an antibody specific for phosphorylated TDP-43 (pTDP-43) sporadic FTD-TDP, C9FTD and control frontal cortex tissue was co-stained for both pTDP-43 and KPNA4. KNPA4 immunolabeling in controls showed a uniform distribution in neurons with both nuclear and cytoplasmic distribution with a weak nuclear TDP-43 labelling (figure 8.5). pTDP-43 inclusions were detectable in both sporadic FTD-TDP and C9-FTD frontal cortex sections but not in controls (figure 8.5). In sporadic FTD-TDP cases and C9-FTD neurons were seen in which KPNA4 was depleted from the nucleus and accumulates in the cytoplasm where it co-localised with pTDP-43 inclusions (figure 8.5). In the FTD-TDP and C9FTD cases investigated in this study the KPNA4 inclusions were more prominent in the C9FTD and always positive for pTDP-43. However numerous cells also showed a nuclear depletion of KPNA4 but no p-TDP-43 inclusion (figure 8.6). Further cases are needed to quantify the severity of KPNA4 inclusions and their association with cytoplasmic pTDP-43 aggregates.

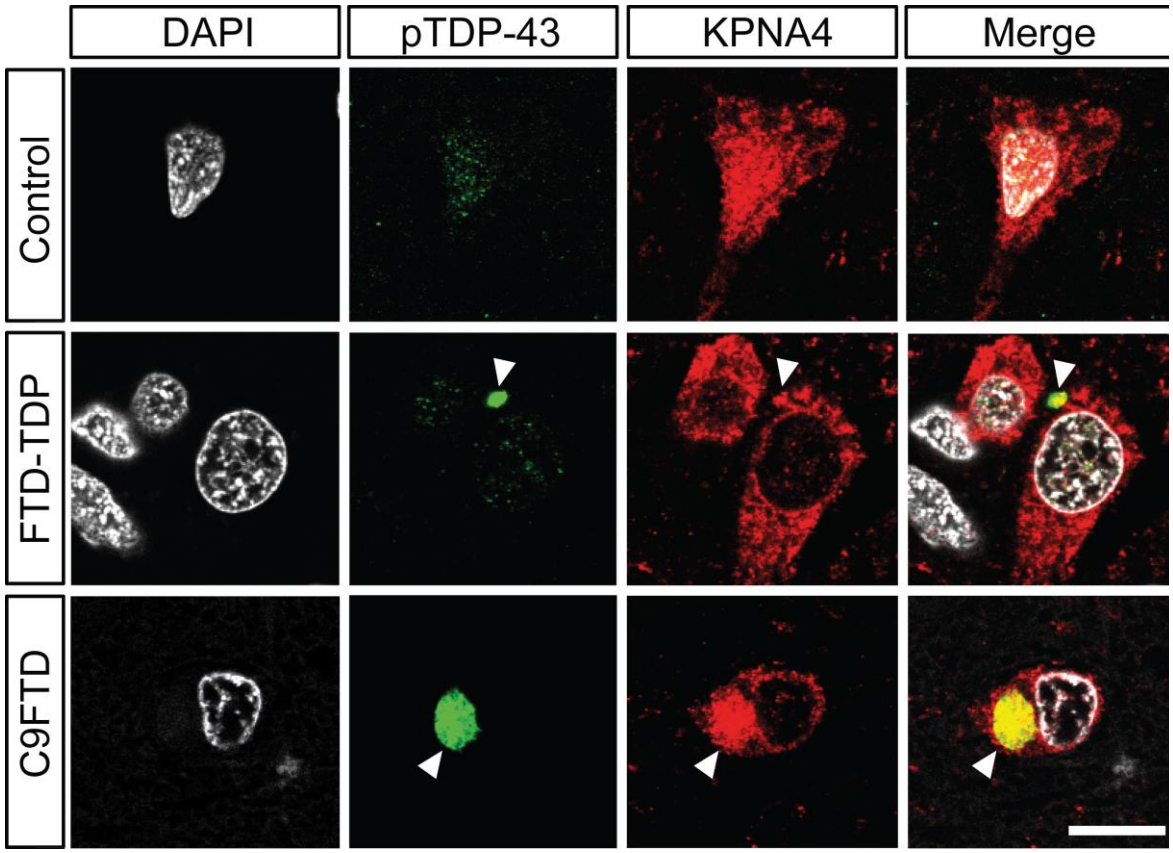

Figure 8.5 | KPNA4 is present in phospho-TDP43 inclusions in sporadic FTD-TDP, and C9FTD frontal cortex. KPNA4 immunolabelling in frontal cortex control cases reveals uniform neuronal distribution in both nucleus and cytoplasm (top row); immunolabeled phospho-TDP-43 inclusions are detectable in both sporadic FTD-TDP (middle row) and C9FTD (lower row). Note that in both sporadic FTD-TDP and C9FTD, but not in controls, immunolabelling reveals nuclear depletion and cytoplasmic accumulation of KPNA4 that overlaps with accumulated phospho-TDP-43 (pTDP-43; arrowheads). Scale bar, 10µm.

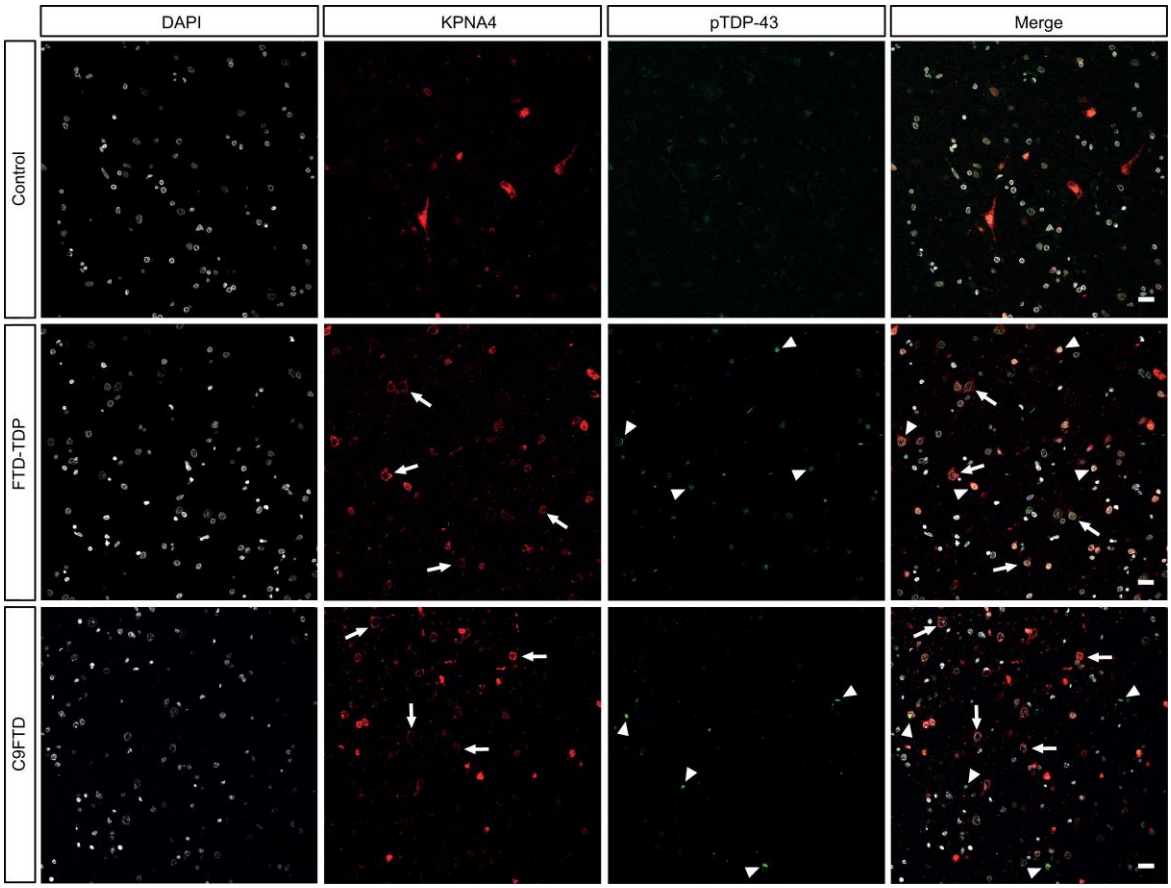

Figure 8.6 | KPNA4 is present in phospho-TDP43 inclusions in sporadic FTD-TDP, and C9FTD frontal cortex. KPNA4 immunolabelling in frontal cortex control cases reveals uniform neuronal distribution in both nucleus and cytoplasm (top row); immunolabeled phospho-TDP-43 inclusions are detectable in both sporadic FTD-TDP (middle row) and C9FTD (lower row). Note that in both sporadic FTD-TDP and C9FTD, but not in controls, immunolabeling reveals nuclear depletion and cytoplasmic accumulation of KPNA4 that overlaps with accumulated phospho-TDP-43 (pTDP-43; arrowheads). Scale bars, 22µm.

8.4 KPNA4 is not sequestered into DPR inclusions in C9FTD frontal cortex

Neuronal cytoplasmic and intranuclear inclusions of RAN translated DPRs are the major and distinctive pathological hallmark of the C9ALS/FTD brain (Mori et al., 2013a; Ash et al., 2013). C9FTD, FTD-TDP and control frontal cortex tissue was co-stained for KPNA4 and each of the three sense DPRs, poly-GP, poly-GA and poly-GR. Poly-GA was stained for using the mouse monoclonal antibody 5E9 antibody validated in Mackenzie et al. (2013) and later confirmed as the best of the currently available poly-GA antibodies for immunohistochemistry (Mackenzie et al., 2015). Star-shaped neuronal cytoplasmic inclusions of poly-GA were seen in C9FTD frontal cortex as has been previously reported (Ash et al., 2013; Mori et al., 2013; Schludi et al., 2015). Cytosolic KPNA4 partially overlapped with poly-GA inclusions (figure 8.7). For poly-GP the previously characterised mouse monoclonal antibody generated in this study was used to probe for the peptide. Staining revealed star like cytoplasmic poly-GP inclusions in the frontal cortex that again partially overlapped with KPNA4 cytoplasmic staining. Poly-GR was stained for using the rat monoclonal 5H9 antibody (Mori et al., 2013), the most sensitive antibody for detecting poly-GR inclusion bodies in patient tissue (Mackenzie et al., 2015). Poly-GR inclusions were relatively rare and difficult to find however when found appeared to also overlap slightly with cytoplasmic KPNA4. However, KPNA4 did not appear to be sequestered into DPR inclusions. Further, as was the case for phospho-TDP-43 staining, cells also displayed KPNA4 nuclear depletion without the presence of a DPR aggregate (figure 8.8). Additional case studies are needed to quantify if nuclear depletion of KPNA4 correlates with any DPR inclusions, in particular poly-GR.

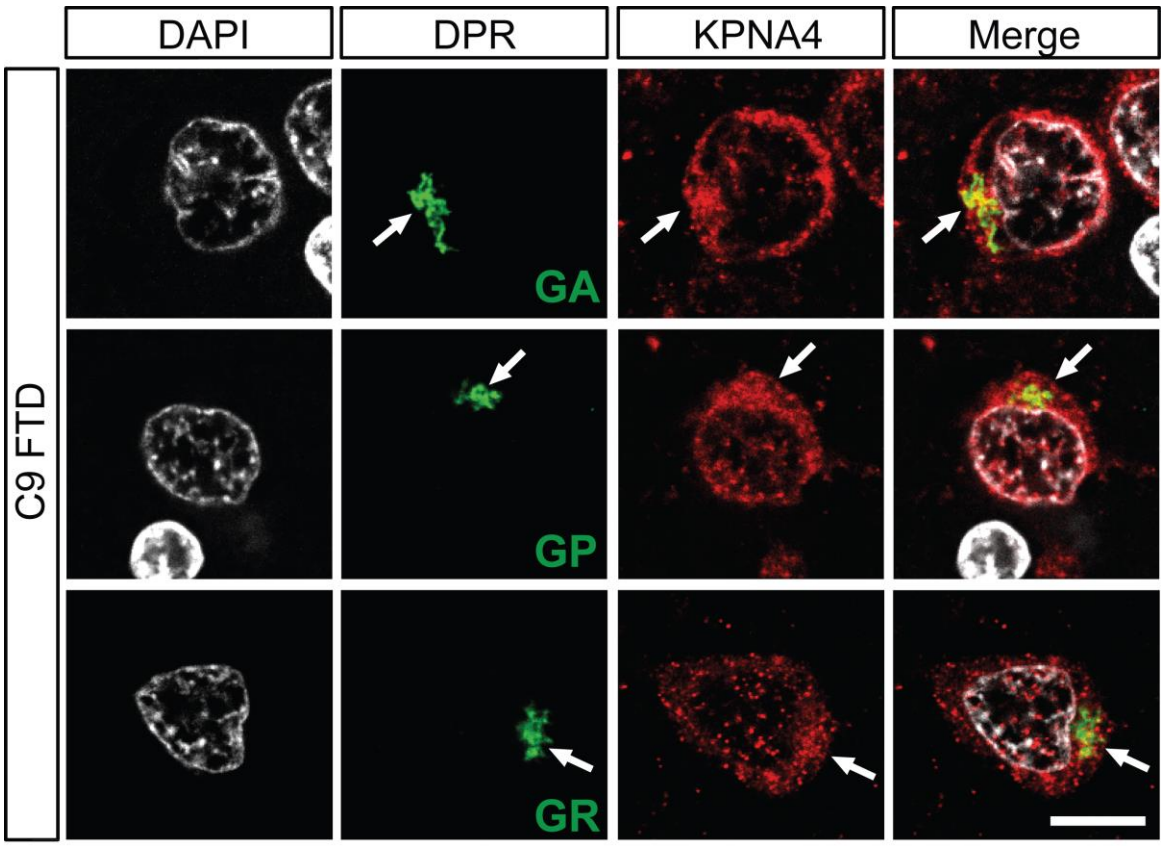

Figure 8.7 | KPNA4 cytoplasmic staining partially overlaps with DPR inclusions. In C9ALS/FTD frontal cortex sections, poly-GA, poly-GP and poly-GR inclusions partially overlap with KPNA4 labelled regions (arrows). Scale bar, 10µm.

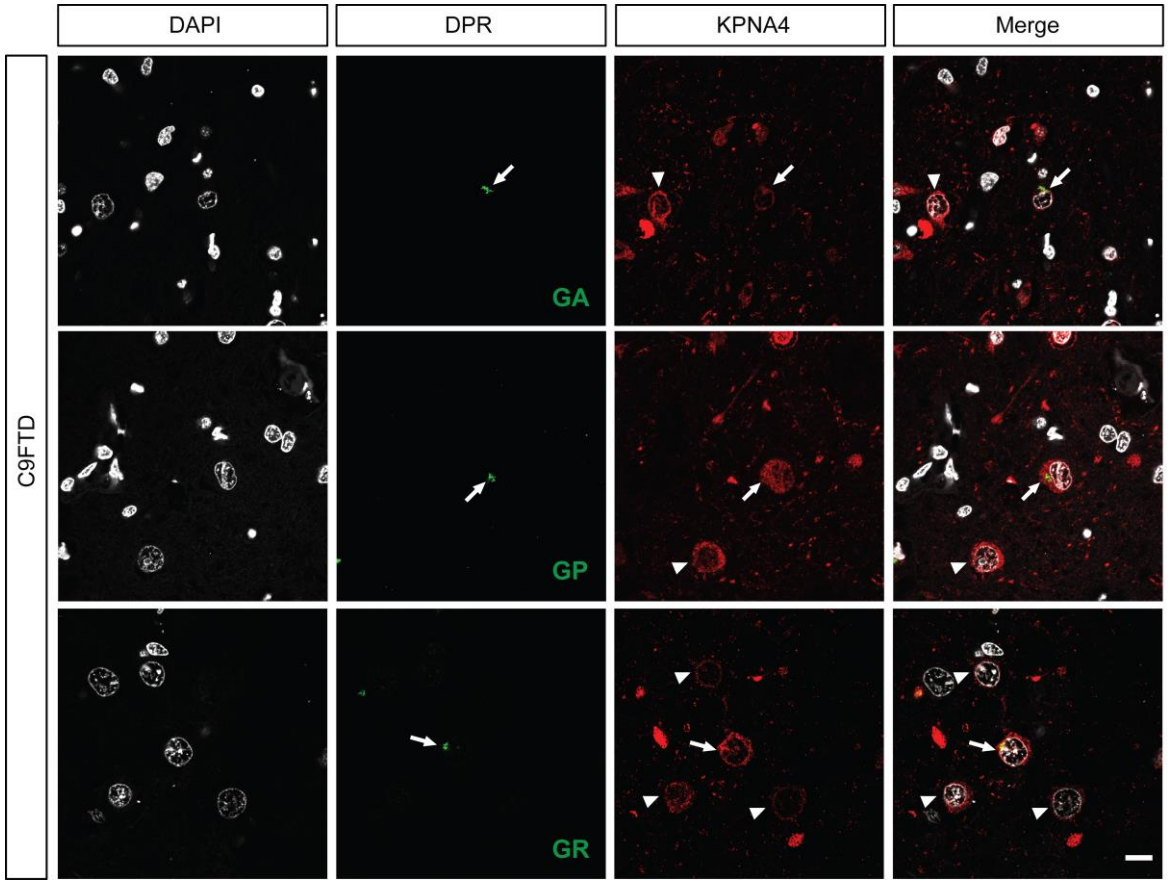

Figure 8.8 | KPNA4 nuclear depletion is seen in cells without DPR inclusions. In C9FTD frontal cortex sections KPNA4 nuclear depletion is seen with (arrows) and without (arrow heads) poly-GA, poly-GP and poly-GR inclusions. Scale bar, 10µm.

8.5 Discussion

8.5.1 KPNA4 mislocalisation is seen in both C9FTD and FTD-TDP frontal cortex

Within the human genome there are seven different KPNA isoforms. The KPNA family of nuclear transport receptors are subdivided into three families named α1, α2 and α3 (Pumroy and Cingolani, 2015). All KPNA isoforms have conserved protein structure consisting of an N-terminal auto inhibitory, importin-β-binding (IBB) domain and a helical C-terminal Arm (Armadillo)-core that binds to protein cargoes with an NLS (Pumroy and Cingolani, 2015). Although the KPNA isoforms have almost identical amino acid sequences and 3D structures – they show a high degree of selectivity for distinct NLS containing protein cargoes (Pumroy and Cingolani, 2015).

The α1 family consists of karyopherin α2 (KPNA2) and the recently discovered karyopherin α7 (KPNA7) (Kelley et al., 2010). KPNA2 is the major NLS binding karyopherin whilst little is known about KPNA7 (Pumroy and Cingolani, 2015). Despite being members of the α1 family, KPNA2 and KPNA7 show distinct subcellular localisations. The steady state cellular localisation of KPNA2 is predominantly cytoplasmic and KPNA7 exclusively nuclear, suggesting divergent functions in the α2 family (Kelley et al., 2010). KPNA2 shows high specificity for several important protein cargos linked to C9ALS/FTD including ADAR3 (Maas and Gommans, 2009), a member of the ADAR RNA editing family of enzymes. ADAR3 has been shown to colocalise with sense RNA foci in C9 IPSC motor neurons and post-mortem motor cortex (Donnelley et al., 2013). Furthermore the protein accumulates in the nucleus of patient tissue and knockdown of ADAR3 decreases RNA foci numbers in iPSC neurons (Donnelley et al., 2013).

The α2 family consists of karyopherin α3 (KPNA3) and karyopherin α4 (KPNA4) (Pumroy and Cingolani, 2015). KPNA3 and KPNA4 have exceptionally high homology in both identity and sequence, with 86% identity and 92% sequence conservation (Pumroy and Cingolani, 2015). KPNA3 and KPNA4 have been shown to be the NTR for several important NLS containing proteins including NF-κB (Fagerlund et al., 2005) and RCC1 (Köhler et al., 1999) (Pumroy and Cingolani, 2015).

The α3 family consists of karyopherin α1 (KPNA1), karyopherin α5 (KPNA5) and karyopherin α6 (KPNA6) (Pumroy and Cingolani, 2015). The family are highly homologous and hence are likely to have similar functions (Pumroy and Cingolani, 2015). The most studied interaction between a protein cargo transported into the nucleus and members of the α3 family is transport of the phosphorylated STAT1 homodimer or the STAT1/STAT2 heterodimer into the nucleus by KPNA1 (Pumroy and Cingolani, 2015).

The KPNAs are involved in the classic nuclear import pathway of which half of all human proteins are transported by (Mihevc et al., 2017). In this pathway, the KPNAs bind to the classic NLS on their protein cargoes. KPNB1 then binds to the KPNA-cargo dimer and the whole complex translocates across the nuclear pore into the nucleus (Mihevc et al., 2017). Once in the nucleus this tripartite complex dissociates by RAN-GTP binding to KPNB1; the KPNB1-RAN-GTP complex is recycled back to the cytoplasm and the KPNAs transfer back into the cytoplasm in a complex of KPNA, RAN-GTP and CAS (cellular apoptosis susceptibility) (Mihevc et al., 2017).

TDP-43 nuclear import occurs through the KPNA/KPNB1 import pathway (Nishimura et al., 2010; Mihevc et al., 2017). KPNA mediated TDP-43 nuclear transport uses the N-terminal bipartite nuclear localisation signal of TDP-43 (Ayala et al., 2008; Winton et al., 2008). Nishimura et al. (2010) using a GST-pulldown found all members of the KPNAs bind to TDP-43 directly hence likely mediate its import into the cytoplasm. However, the GST-pulldown showed TDP-43 has preferential binding to members of the α2 family – KPNA3 and KPNA4. Additionally, KPNA binding to TDP-43 was abolished using mutant TDP-43ΔNLS. Although RNAi interference to all karyopherin α isoforms did not induce a cytoplasmic accumulation of TDP-43 suggesting redundancy in KPNA isoform transport of TDP-43. Interestingly Nishimura et al. (2010) saw RNAi knockdown of two members of the karyopherin β family - karyopherin-β1 (KPNB1) and CAS did lead to the cytoplasmic accumulation and formation of TDP-43 aggregates although GST pulldown revealed KPNB1 binding to TDP-43 is much weaker than TDP-43 binding to KPNAs 3 and 4. Nishimura et al. (2010) found in FTD-TDP post-mortem cortical lysate that CAS and KPNA2 were significantly downregulated. Immunohistochemical analysis revealed nuclear KPNA2 and both CAS nuclear and cytoplasmic immunoreactivity were reduced in patient cortical neurons. In sporadic ALS spinal cord lysate KPNA2 and

KPNA6 protein levels were significantly increased with KPNA2 cytoplasmic staining increased in ALS spinal cords compared to controls. Hence the classic nuclear import pathway is affected in both diseases characterised by TDP-43 cytoplasmic accumulation and aggregation. Intriguingly CAS was strongly affected in FTD-TDP but not sporadic ALS suggesting differences between the two diseases. Interestingly KPNB1 knockdown caused a nuclear loss of KPNAs – whilst CAS knockdown caused KPNA cytoplasmic accumulation; furthermore, as noted previously KPNB1 and CAS knockdown cause cytoplasmic TDP-43 accumulation. Nishimura et al. (2010) propose that this loss of KPNAs from the cytoplasm or an inhibition of nuclear import could lead to increased cytosolic TDP-43.

In the current study both a nuclear loss and nuclear accumulation of KPNA4 is seen in sporadic FTD-TDP and C9-FTD. This finding taken together with the Nishimura et al, (2010) results indicates a disruption of the KPNA/KPNB1 import pathway may underlie disrupted transport of TDP-43 and its subsequent cytoplasmic accumulation. Indeed, other important members of the KPNA/KPNB1 import pathway have demonstrated to be affected in disease associated tissue. For example, nuclear contour irregularity and a strong reduction in nuclear KPNB1 was seen in ALS spinal cords (Kinoshita et al., 2009). Xiao et al. (2015a) demonstrated that KPNB1, and Ran-GTPase pathology correlates with cytoplasmic TDP-43 accumulation and aggregation. Interestingly a loss of the C9ORF72 short isoform from the nuclear membrane also correlated with KPNB1, RAN and TDP-43 mislocalistion. This supports the Nishimura et al. (2010) cell work showing KPNB1 disruption causes enhanced cytosolic TDP-43 localisation and work in a progranulin knockout mouse model which demonstrated RAN-GTPase nuclear trafficking correlated with a reduction in TDP-43 from the nucleus (Ward et al., 2014). Not all DAPI labelled cells show KPNA4 immunolabeling. There are 7 KPNA human isoforms, although they have different cargoes, there is likely redundancy in their function. Indeed all 7 have been shown to bind TDP-43 to different extents (Nishimura et al., 2010). Hence the levels of individual KPNA isoforms could depend on the precise needs of a cell at a certain time. For example, both transcript and protein levels of the KPNAs dynamically change during oocyte development in the mouse (Mihalas et al., 2015). Whether the transcript and protein levels of KPNAs are also dynamic in adult cells, such as neurons, is unknown, however it is plausible to

speculate this could occur as different cargoes are transported by different KPNAs. There could also be cell-type specific expression differences of KPNA isoforms between neurons, different neuronal subtypes and glia. For example, the different KPNA isoforms are differentially expressed in oligodendrocytes progenitor cells (Latiman et al., 2017). Co-staining of KPNA4 with markers for neurons, astrocytes and oligodendrocytes should be performed to investigate this possibility.

The data from the present study supports pathological disturbances in members of the classical nuclear import pathway leads to accumulation of TDP-43 in the cytoplasm. A loss of KPNAs from the cytoplasm, or the inhibition of their import into the nucleus, may lead to cytoplasmic accumulation of TDP-43 as has been proposed by Nishimura et al, (2010). Further work is needed to see whether the KPNA4 pathology seen in current body of work correlates with cytoplasmic TDP-43 accumulation however its sequestration into TDP-43 cytoplasmic inclusions suggests this may be the case.

The majority of disease causing proteins in ALS and FTD are RNA binding proteins, in disease these proteins are depleted from the nucleus and form inclusions in the cytoplasm. How this nuclear depletion and subsequent cytoplasmic accumulation initiates is unknown particularly in sporadic forms of both diseases (Boeynaems et al., 2016b). Boeynaems et al. (2016b) through both a review of the literature and bioinformatics analysis found RNA binding proteins and nucleocytoplasmic transport have major roles in ALS and FTD pathogenesis. The authors propose defective nuclear transport bridges nuclear loss and cytoplasmic accumulation, having a key initiating role in disease. In support of this studies have shown without doubt nucleocytoplasmic transport is a implicated by both G_4C_2 RNA and DPR toxicity (Zhang et al., 2015, Freibaum et al., 2015, Jovičić et al., 2015, Boeynaems et al., 2016). The observation that KPNA4 is perturbed sporadic FTD-TDP cases further supports this notion that disturbances in nucleocytoplasmic transport are a key disease mechanism in ALS and FTD.

There are direct mechanistic links between pathological gain of function features of the *C9ORF72* repeat expansion and impaired nucleocytoplasmic transport. Shi et al., 2017) found poly-PR shifts nuclear pore protein FG repeats toward a polymerised state proposing this causes impairments in nuclear cytoplasmic shuttling. Defective

nuclear pore functioning could lead to a disruption in KPNA shuttling and subsequent TDP-43 mislocalisation and aggregation as proposed by Nishimura et al. (2010). KPNA4 nuclear depletion was observed in cells however which did not contain a DPR inclusion; however it could be the soluble forms of these DPRs that are the toxic species responsible which would not be detected using current antibodies (Gitler and Tsuiji, 2016). KPNA4 mislocalisation and aggregation however was also seen in *C9ORF72* negative, sporadic FTD-TDP post mortem tissue. The DPRs are obviously not present in *C9ORF72* negative sporadic cases; hence the question is how KPNA4 is affected in *C9ORF72* negative forms of FTD. A recent study by Woerner et al. (2016) suggests defective proteostasis could impair nuclear import. This study found both cytoplasmic β protein aggregates and human disease proteins such as TDP-43 caused mislocalisation of aggregation of nuclear pore proteins and nuclear import receptors which led to impaired protein import and mRNA export. For example Woerner et al. (2016) investigated the transport of NF-κB. NF-κB is a transcription factor transported into the nucleus by KPNA3 and KPNA4. In cells expressing cytoplasmic β protein aggregates and TDP-43, NF-Kb nuclear import was impaired – the protein remained cytoplasmic. Although the authors did not co-stain TDP-43 and KPNA4 in their cellular model, cytoplasmic β protein aggregates were shown to sequester KPNA4 into inclusions. The Woerner et al. (2016) paper shows a clear dysfunction and co-aggregation of KPNA4 because of cytoplasmic protein aggregates; similar results were seen for KPNA2. If as this paper suggests nuclear transport deficits are caused by cytoplasmic protein aggregation; defective nuclear transport would be predicted to be seen in many neurodegenerative diseases characterised by cytoplasmic proteinaceous inclusions – and not just specific to the *C9ORF72* expansion. Woerner et al. (2016) saw nuclear pore and KPNA dysfunction following expression Htt96Q and ParkinΔC. RanGAP1, Nup88 and Nup62 aggregate with mutant huntington in two in HD mouse models (Grima et al., 2017). RanGAP1 aggregated and Nup62 was mislocalised to either the cytoplasm or the nucleus in HD striatum (Grima et al., 2017). Hence defective proteostasis leading to the aggregation of TDP-43 will disrupt nuclear import and export by sequestering nucleocytoplasmic transport proteins that interact with TDP-43 – including direct bindings partners like KPNA4. This sequence of events would suggest defective nucleocytoplasmic transport occurs after and rather is caused by cytoplasmic TDP-43 accumulation. The nuclear depletion of KPNA4 was seen in the sporadic FTD-

TDP frontal cortex was not however always accompanied by a TDP-43 aggregate. Hence aggregation of TDP-43 may not be the sole cause of the KPNA4 mislocalisation observed. TDP-43 regulates the splicing and transcription of nucleocytoplasmic transport proteins (Dormann and Haass, 2011). Hence a loss of TDP-43 function via cytoplasmic accumulation (without aggregation) and/or nuclear depletion may also perturb nucleocytoplasmic transport. How nucleocytoplasmic transport and TDP-43 may be disturbed in non-*C9ORF72* ALS/FTD is discussed in greater depth in the final discussion of this thesis.

KPNA4 has many interesting properties and functions that makes its dysfunction interesting and worth studying further. KPNA4, KPNB1, RanBP1 via a mechanism that involves intra-axonal protein synthesis work together to signal nerve injury to induce regeneration of an injured axon in the PNS (Twiss et al., 2016). KPNA4 is required for oligodendrocyte differentiation (Latiman et al., 2017). Of particular interest is the specific role of KPNA4 in the shuttling of NF-Kb from the nucleus to the cytoplasm. NF-Kb is a protein complex that regulates a proinflammatory signalling, the pathway is activated by proinflammatory cytokines such as IL-1 and TNFα (Lawrence, 2009). Upon activation the NF-κB complex translocates into the nucleus where it acts as a transcription factor inducing the expression proinflammatory genes such as cytokines, chemokines, and adhesion molecules (Lawrence., 2009). The NF-kB family of transcription factors are p65, p50, p52, c-REL, and RELB, these proteins all form homodimers and heterodimers with each other during NF-Kb inflammatory signalling (Roussos et al., 2013). In neurons, the most common NF-Kb heterodimer is the p65 and p50 complex (Gutierrez and Davies, 2011). A classic hallmark of familial and sporadic ALS patients is neuroinflammation including astrogliosis and microglial activation (Frakes et al., 2014). Loss of function mutations in optineurin have been found in ALS patients (Maruyama et al., 2010); optoneurin knockdown has been shown to cause neurodegeneration by NF-κB activation (Akizuki et al., 2013). Another ALS causing gene *TBK1* is also linked to inflammation (Oakes et al., 2017).

NF-Kb p65 mRNA and protein levels are high in sporadic ALS spinal cords compared to controls and TDP-43 co-localises with p65 in both patient glia and neurons and TDP-43 transgenic mice (Swarup et al., 2011). TDP-43 also acted as co-activator of p65 as glial cells expressing higher levels of TDP-43 had an

increased inflammatory response. Inhibition of NF-Kb activity reduced neurodegeneration and motor phenotypes in TDP-43 transgenic mice (Swarup et al., 2011). Hence TDP-43 toxicity may in part be mediated by a disturbance of the NF-Kb signalling pathway. Both immunoprecipitation and immunofluorescence demonstrated an interaction between TDP-43 and p65 in temporal lobe neuronal nuclei in patients with mild cognitive impairment (Ohta et al., 2014). It has recently been shown that TDP-43 prevents TNF-α induced NF-κB nuclear translocation inhibiting its activity (Zhu et al., 2015). This TDP-43 mediated inhibition occurred through TDP-43 competing for binding with KPNA4 as when TDP-43 was overexpressed p65 association with KPNA4 was significantly decreased in TNF-α treated cells compared to non-treated control cells (Zhu et al., 2015). As noted previously GST-pull down by Nishimura et al. (2010) showed KPNA4 to be the strongest binder of TDP-43. Zhu et al. (2015) propose two hypotheses as to how, TDP-43 may compete with p65 for KPNA4 binding preventing its nuclear localisation thereby inhibiting its activity. Firstly TDP-43 could compete with p65 for KPNA4 binding in the cytoplasm which reduces p65 nuclear transport; or secondly TDP-43 binds to KPNA4 in the nucleus reducing its cytoplasmic levels resulting in defective p65 nuclear translocation. Additionally, NF-κB gene and protein levels and its nuclear activation are significantly reduced in the superior temporal gyrus of schizophrenia patients – this was associated with a decrease in the levels of KPNA4 and an SNP within the KPNA4 locus (Roussos et al., 2013). This SNP in KPNA4 was associated with increased schizophrenia susceptibility (Roussos et al., 2013); interestingly repeat expansions in *C9ORF72* are causes of, albeit rare, schizophrenia spectrum disorders (Galimberti et al., 2014).

8.5.2 Conclusion

The data from the present study further adds to the increasing body of evidence showing nucleocytoplasmic transport is disrupted in *C9ORF72* expansion associated disease. Preliminary evidence from a small number of cases has shown KPNA4 was depleted from the nucleus in sporadic FTD-TDP frontal cortex neurons. Further work is needed to establish whether KPNA4 nuclear depletion correlates with a cytoplasmic accumulation of wild-type TDP-43, TDP-43 inclusions or DPR pathology

in larger number of cases. It will also be important to investigate KPNA4 localisation in C9ALS and SALS in spinal cord neurons.

Chapter 9: General discussion

This study has investigated G_4C_2 derived DPR protein toxicity in a pure G_4C_2 repeat fly model and a DPR protein only fly model. Both models converge to show the sense DPR poly-GR is highly neurotoxic; poly-GA and poly-GP were associated with milder toxicity in pure repeat flies whilst the G_4C_2 RNA did not appear to confer toxicity. C9ALS/FTD, like 97 % of ALS and 45 % of FTD cases is a TDP-43 proteinopathy, characterised by the nuclear depletion and formation of cytoplasmic aggregates of the RNA binding protein TDP-43 (Ling et al., 2013). TDP-43 pathology correlates with neurodegeneration and clinical symptoms in C9ALS/FTD (Mackenzie et al., 2014) hence TDP-43 dysfunction is thought to be the most likely effector of neuronal loss in C9ALS/FTD (Edbauer and Haass, 2016). A key question in C9ALS/FTD is the link between the hexanucleotide repeat expansion and TDP-43 pathology. Edbauer and Haass, (2016) have proposed an amyloid cascade like hypothesis for C9ALS/FTD pathogenesis in which specific stressors cause the deposition of TDP-43 into the cytoplasm and its subsequent aggregation. The data from the present study demonstrate that the DPRs are key to this disease cascade. When an alternative codon poly-GR construct is expressed in flies a striking cytoplasmic accumulation of the Drosophila homologue of TDP-43 – *TBPH* – is seen. In addition Poly-GA expression led to the formation of cytoplasmic inclusions of *TBPH* whereas no changes in TBPH localisation were observed following expression of the G_4C_2 RNA only. These results strongly suggest the DPRs are involved in the development of TDP-43 pathology seen in C9ALS/FTD. These alterations were accompanied by changes in the localisation of importins α2 and α3, the *Drosophila* homologues of karyopherins KPNA2 and KPNA4, respectively. However, nucleocytoplasmic transport components such as RanGAP and nuclear pore complex proteins previously implicated in C9ALS/FTD were morphologically unperturbed. No changes in the localisation of nucleocytoplasmic transport proteins were seen following the expression of the G_4C_2 RNA only. Further, the cytoplasmic mislocalisation of *TBPH* enhanced DPR levels and cytotoxicity. Similar phenotypes were observed for patient frontal cortex in sporadic FTD-TDP and C9FTD; with KPNA4 being depleted from the nucleus and overlapping with TDP-43 inclusions.

Taken together, these findings establish DPR accumulation as a cause of TDP-43 proteinopathy.

9.1 DPR levels and identity correlate with toxicity in a G_4C_2 repeat *Drosophila* model

RAN translation and motor impairment was investigated in multiple transgenic flies containing different lengths of the G_4C_2 repeat sequence located in the 3' untranslated region of the red fluorescent protein DsRed2. These constructs were designed based upon a similar approach used by Li et al. (2008) and Yu et al. (2011) to model CAG and CUG RNA expansions in *Drosophila* for spinocerebellar ataxia type 3 (SCA3) and Myotonic Dystrophy Type 1 (DM1). DsRed2 was chosen as it has no connection with disease pathogenesis and did not confound repeat associated toxicity in these two studies. The DPR expression profile was investigated in flies harbouring repeat lengths of 8, 32, 38, 56 and 64. No DPRs were detectable in flies with <32 repeats indicating a length dependent threshold for RAN translation. The levels and identity of the DPRs produced differed in a construct dependent manner which correlated with the linker after the DsRed2 STOP codon 5' of the start of the repeats as flies harbouring constructs with the same 5' linker sequence produced the same DPR expression profile. This differences in DPR expression profile across the different constructs facilitated the investigation of the toxicity mediated by different levels and combinations of DPRs.

The 32 and 64 repeat flies both produced high levels of poly-GP and poly-GA compared to the other repeat lines. Both the 32 and 64 repeat flies had a late onset motor phenotype characterised by a reduction in climbing performance and activity at day 40 compared to all other genotypes tested at this time point. Of interest is the comparison of these two lines with the 56 repeat flies. The 56 repeat flies produced much lower levels of poly-GP and poly-GA and displayed no signs of motor impairment at day 40 in the assays used in this study. Importantly transgene expression levels (as measured by DsRed2 protein levels) were similar between the 56 repeat flies and the 32 and 64 repeat flies suggesting the phenotypes seen in the 32 and 64 repeat lines is not attributable to G_4C_2 RNA toxicity. This supports work in other fly models that have demonstrated expression of the G_4C_2 RNA on its own

(without DPRs) is not toxic. Rather this data suggests the behavioural impairment seen in the 32 and 64 repeat lines is caused by the high levels of poly-GP and/or poly-GA. Hence poly-GP and/or poly-GA is toxic at high enough levels in aged neurons. This study was unable to distinguish whether this age related late onset phenotype was caused by poly-GP or poly-GA (or a combination of the two). However, evidence for poly-GP toxicity in other model systems is mild at best (Freibaum & Taylor, 2017) whereas there is a general consensus that poly-GA is toxic when expressed at high enough levels (Moens et al., 2017, Freibaum & Taylor, 2017) though UPS dysfunction and ER stress (May et al., 2014; Zhang et al., 2014; Zhang et al., 2016). Further poly-GP expression was higher than poly-GA in the 56 line indicating the very low levels of poly-GA may be why no day 40 phenotype was seen in these flies; conversely this suggests high levels of poly-GA expression may be the biggest contributor to the day 40 motor impairment observed in the 32 and 64 repeat lines. Indeed Mizielinska et al. (2014) found poly-GA causes a reduction in life span from day 50 onwards in flies. Hence investigations into how poly-GA contributes to age-related toxicity in flies is an interesting future avenue of research.

A severe early onset motor impairment was observed in the 38 repeat flies that produced all three sense DPRs – poly-GP, poly-GA, poly-GR. This motor phenotype was again attributable to DPR toxicity rather than G_4C_2 RNA toxicity as even though transcript levels were higher in the 38 repeat line (as evidenced by DsRed2 expression levels), the generation of a fly line with 2x copies of the 64 construct was used to confirm this did not contribute to the early phenotype. Flies harbouring two copies of the 64 repeat construct had significantly higher levels of DsRed2 protein compared to both flies with a single copy of the 64 construct and the 38 repeat line. There was no significant difference in day 5 climbing performance between the 2x 64 flies and the 1 x 64 flies, however both were significantly impaired compared to the 38 repeat line (appendix 2; figure A2). Hence DPR toxicity (and not DsRed2 expression or G_4C_2 RNA levels) causes the early-day 5 behavioural phenotype seen in the 38 G_4C_2 repeat line. The 38 repeat line is unique as it was the only fly line in which poly-GR was detected; hence the presence of poly-GR correlates with severe early onset toxicity. This is in line with other studies in *Drosophila* that poly-GR highly neurotoxic (Mizielinska et al., 2014; Yang et al., 2015; Freibaum et al., 2016).

In summary this study demonstrates the DPRs and not the G_4C_2 RNA correlates with toxicity in a *Drosophila* model of the G_4C_2 repeat expansion in *C9ORF72*, consistent with what has been observed in other Drosophila models (Mizielinska et al., 2014; Tran et al., 2015). However, the extent of DPR toxicity was dependent upon DPR identity and expression levels. Both poly-GP and poly-GA appear well tolerated, however when the expression levels are high enough in combination with age they appear sufficient to trigger late onset disease related phenotypes. Poly-GR expression on the other hand correlated with a severe early onset motor impairment. Flies expressing high levels of poly-GA and poly-GP (not poly-GR) had no behavioural impairment at days 5 and 20 whereas by day 20 flies producing detectable levels of poly-GR were unable to climb. This demonstrates poly-GR is highly toxic compared to poly-GA and poly-GP in line with the general consensus in the field that the arginine DPRs specifically have an important pathogenic role in C9ALS/FTD (Gitler and Tsuiji, 2016; Freibaum and Taylor, 2017). The lack of spatial correlation between the DPRs and regions of the brain showing the most severe neurodegeneration is still a contentious issue (Gitler and Tsuiji, 2016). However post-mortem analysis only visualises neurons that have survived; if the arginine DPRs are as toxic as they appear to be in model systems these neurons may already be lost. Further soluble DPR species, which cannot be visualised with current antibodies, rather than aggregates, may mediate toxicity. For example, aggregates of amyloid-beta do not correlate with neurodegeneration in Alzheimer's disease; rather it is thought the soluble oligomeric form of the protein is what confers toxicity (Edbauer and Haass, 2016).

9.2 DPR, but not G_4C_2 RNA expression cause *Drosophila* KPNA2, KPNA4 and TDP-43 mislocalisation

C9ALS/FTD is TDP-43 proteinopathy characterised by the nuclear depletion of TDP-43 and its accumulation into aggregates in the cytosol (Mackenzie et al., 2015). TDP-43 pathology correlates with neurodegeneration in C9ALS/FTD (Mackenzie et al., 2015) hence it is thought TDP-43 is the most likely effector of neuronal death in C9ALS/FTD (Edbauer and Haas, 2011). The question is how does TDP-43 become depleted from the nucleus in expansion carriers. Murine models show TDP-43 aggregation occurs after the formation of both DPRs and RNA foci (Liu et al 2016a).

Furthermore, post-mortem analyses on *C9ORF72* expansion carriers who died early of the disease show abundant DPR pathology only suggesting DPR pathology precedes that of TDP-43 (Proudfoot et al., 2015; Baborie et al., 2015; Vatsavayai et al., 2016). Based on such findings an amyloid cascade hypothesis for C9 mediated neurodegeneration has been proposed in which stressors causes TDP-43 aggregation and cytoplasmic mislocalisation (Edbauer and Haass, 2016). The DPRs and/or RNA foci are argued to be such stressors (Edbauer and Haass, 2016). Hence this project investigated whether the G_4C_2 RNA and the sense DPRs, poly-GR and poly-GA, are on their own capable of causing TDP-43 mislocalisation.

The expression of the G_4C_2 RNA from a construct with forms RNA foci but does not produce DPRs (Mizielinska et al., 2014) in *Drosophila* salivary gland cells caused no changes in the normally nuclear localisation of Drosophila TDP-43 (*TBPH*). This finding suggests the G_4C_2 RNA is not sufficient for TDP-43 mislocalisation in C9ALS/FTD. Indeed, this supports findings from other Drosophila models of the repeat expansion showing G_4C_2 RNA expression without the downstream of production of DRPs is non-toxic (Mizielinska et al., 2014; Tran et al., 2015). Poly-GR expression however on its own caused a cytoplasmic accumulation of *TBPH*, whereas poly-GA caused the formation of inclusions of *TBPH* in the cytoplasm that co-localised with poly-GA. These results indicate the DPRs, and not the G_4C_2 RNA, are the key stressor in this cascade, accumulating early during the disease progression and subsequently causing TDP-43 to be mislocalised to the cytoplasm. Furthermore, it was found cytoplasmic *TBPH* expression enhances G_4C_2 related motor phenotypes indicating DPR-mediated mislocalisation of TDP-43 contributes to toxicity in C9ALS/FTD. Interestingly forcing *Drosophila* TDP-43 into the cytoplasm was shown to enhance DPR production in G_4C_2 repeat expressing flies, suggesting this enhanced phenotype could be attributable to increased DPR levels. Not only does this further support the findings from the pure repeat model investigated in this study that correlated phenotypes with DPR toxicity rather than G_4C_2 RNA, but also indicates some kind of feedback mechanism where DPR mediated cytoplasmic TDP-43 accumulation enhances DPR production; potentially further exacerbating TDP-43 mislocalisation and dysfunction.

Dormann and Haass, (2011) propose a multi-hit hypothesis for TDP-43 proteinopathy in which defective nuclear import is the first stage in the initiation of TDP-43 pathology. Numerous screens in flies and yeast have shown poly-GR and poly-PR directly bind proteins involved in nuclear import and export; further modulating nuclear transport can supress and enhance G_4C_2 RNA and arginine DPR toxicity (Zhang et al., 2015; Freibaum et al., 2015; Jovičić et al., 2015; Boeynaems et al., 2016; Lee et al., 2016). Furthermore, the localisation and function of nucleocytoplasmic transport proteins have been shown to be disrupted in *in vivo* models of DPR and G_4C_2 RNA toxicity and in the C9ALS/FTD human brain (Zhang et al., 2015; Jovičić et al., 2015; Zhang et al., 2016). TDP-43 is shuttled into the nucleus via the classical KPNA/KPNB nuclear import pathway. In this pathway, the KPNAs bind to the classic NLS on their protein cargoes. KPNB1 binds the KPNA-cargo dimer, the complex then translocates across the nuclear pore into the nucleus (Mihevc et al., 2017). Once in the nucleus this tripartite complex dissociates by RAN-GTP binding to KPNB1; the KPNB1-RAN-GTP complex is recycled back to the cytosol and the KPNAs transfer back into the cytoplasm in a complex of KPNA, RAN-GTP and CAS (Mihevc et al., 2017). To explore potential mechanisms as to how the localisation of *Drosophila* TBPH might be disrupted by poly-GR and poly-GA expression the morphology of proteins involved in shuttling TDP-43 into the nucleus from the cytoplasm was investigated. No abnormities in the NPC (as evidenced by staining for FG-repeat nups using MAB414 and Nup50 antibodies) or proteins involved in the Ran pathway (as evidenced by staining for the *Drosophila* homologues of RanGAP1 and RCC1) were observed. Although the *Drosophila* homologue of RCC1 was present in some poly-GA inclusions; however no nuclear depletion of RCC1 was reported as previously seen in C9ALS iPSCs (Jovičić et al., 2015). Furthermore, expression of the G_4C_2 RNA also did not disrupt the morphology of any of the nucleocytoplasmic transport proteins investigated arguing against the repeat RNA being the primary toxic species in C9ALS/FTD. This is in line with the work on the DsRed2-G_4C_2 repeat fly model characterised in this thesis in which toxicity correlated with DPR identity and expression levels.

However, a striking nuclear reduction of the *Drosophila* homologues of KPNA2 (*Pendulin*) and KPNA4 (*Importin-α3*) was observed following poly-GR expression. Poly-GR could cause a nuclear depletion of KPNA proteins through several

mechanisms. Firstly, it has been shown poly-GR causes oxidative stress (Lopez-Gonzalez et al., 2016), the KPNAs are known to be are mislocalised in cells exposed to oxidative stress (Kodiha et al., 2008). Poly-GR could induce deficits in the nuclear pore complex (NPC) leading to KPNAs mislocalisation. No morphological deficits were seen in the NPC following poly-GR expression in salivary gland cells however this does not necessarily mean the function of the NPC is not impaired. Aggregates of nuclear pore proteins such as those observed in patient tissue (Zhang et al., 2015) may take time to occur and are only detectable at end-stage. The larval salivary glands have large nuclei which allow a detailed analysis of the morphology of the nuclear pore; however, they represent a very early time point in the *Drosophila* life-cycle. NPC morphology could be studied in the neurons of aged poly-GR flies using the GAL80^{ts} or GeneSwitch systems to temporally control protein expression. The arginine DPRs perturb liquid-liquid phase separations (LLPS) important for the function of the NPC (Frey and Görlich, 2007; Hülsmann et al., 2012; Lee et al., 2016; Schmidt and Görlich, 2016). The formation of the nuclear pore involves a high local concentration of FG-repeat domain containing proteins which imposes pore structure; the FG-repeat domain is a type of LCD facilitating phase transitions in the pore (Frey and Görlich, 2007; Hülsmann et al., 2012; Schmidt and Görlich, 2016). Phase separations of the FG-repeat domains into dense polymer networks gives the NPC selectivity as the resulting FG hydrogels act as barriers to normal macromolecules but are permeable to nuclear transport receptors which are able to bind to the FG repeats in addition to their transport cargoes (Schmidt and Görlich, 2016) Shi et al. (2017) found poly-PR binds to FG-repeats stabilising their fibrillar form; this is predicted to reduce the dynamics of the central channel thereby disrupting the transport of molecules across the pore (Kim and Taylor, 2017). Poly-GR may also block the nuclear pore like poly-PR meaning the KPNAs are unable to diffuse through the NPC, therefore becoming depleted from the nucleus and accumulating in the cytoplasm. One experiment to test whether NPC dysfunction may contribute would be to overexpress NPC proteins to see if this rescues the KPNA nuclear depletion seen following poly-GR64 expression.

Poly-GR also disrupts stress granule (SG) dynamics through perturbed LLPS. In fact, nucleocytoplasmic transport proteins have been shown to be components of SGs. Chang and Tarn, (2009) found both transportin-1 and KPNB1 associate with

SGs following arsenite treatment, furthermore transportin-1 was involved in the movement of macromolecules between SGs and P-bodies. Both KPNA2 and RAN are also found abundantly in SGs following arsenite treatment (Fujimura et al., 2010). Mahboubi et al. (2013) expanded this further finding all three of the KPNA subfamilies (α1, α2 and α3) are components of SGs under stress conditions. The three KPNA subfamilies were recruited to SGs by several different stressors including oxidant diethyl maleate, arsenite and heat shock treatment. This recruitment appeared specific to SGs as the KPNAs were not found in P-bodies. Mahboubi et al. (2013) also replicated the results of Chang and Tarn. (2009) showing KPNB1, which binds the KPNAs in the cytoplasm, is recruited to SGs however this was not the case for CAS which binds the KPNAs in the nucleus. The KPNAs and KPNB1 were present in more than 90% of stress granules in total. The presence of these nuclear transport receptors in SGs has led to the suggestion that they are also involved in the assembly of SGs (Mahboubi et al., 2013). The mechanism by which nuclear import factors are recruited to stress granules is not yet known. Nuclear transport proteins may be involved in the assembly of SGs by delivering SG components to the growing granule (Mahboudi et al., 2013); indeed, KPNA2 knockdown causes a reduction in the size of SGs (Fujimura et al., 2010). The interaction between nucleocytoplasmic transport proteins and SGs may be mediated by their mRNA interactions. Mahboubi et al. (2013) found KPNA2 associates with poly(A) mRNA, however it is also easy to speculate that they may be recruited to SGs by their protein cargoes, such as TDP-43, a known SG component. It will therefore be important to look at SGs in poly-GR expressing salivary glands to see if KPNA and KPNB proteins are present in the poorly dynamic SGs induced by poly-GR (Lee et al., 2016). Modulating the levels of SG proteins could alleviate KPNA mislocalisation. Altering stress granule dynamics to promote their disassembly may rescue poly-GR induced cytoplasmic accumulation of *TBPH*. Indeed Lee et al. (2016) found knockdown of several SG proteins; including the *Drosophila* homologue of ataxin-2, completely rescues poly-GR and poly-PR survival. It would be interesting to see if this knockout of ataxin-2 could rescue *TBPH* accumulation in the cytoplasm in poly-GR64 flies. It is also important to stain for SG markers to see if they co-localise with cytoplasmic *TBPH* in poly-GR64 flies.

Poly-GR64 expression in salivary gland cells and the L3 larval eye disc caused *Drosophila* TDP-43 to accumulate in the cytosol. Could the KPNA dysfunction observed in poly-GR expressing salivary gland cells possibly mediate this? Both KPNA2 and KPNA4 bind to TDP-43 in the cytoplasm and are involved in its shuttling into the nucleus (Nishimura et al., 2010). Nishimura et al. (2010) found a knockdown of the individual KPNAs (including KPNA2 and KPNA4) does not result in cytosolic TDP-43 accumulation; although the authors suggest that this may be because TDP-43 is bound and transported into the nucleus by all the KPNAs. However, when KPNB proteins were knocked down in cells TDP-43 was found to accumulate in the cytoplasm. Importantly knockdown of the KPNB proteins KPNB1 and CAS disrupted KPNA localisation. Nishimura et al. (2010) found CAS was markedly downregulated in FTD-TDP frontal cortex and consistent with the western blot data, CAS immunoreactivity was reduced in the cytoplasm and nucleus in FTD-TDP frontal cortex compared to controls. This led Nishimura et al. (2010) to propose inhibition of nuclear import of KPNAs leads to cytosolic TDP-43 accumulation (due to their direct interaction with the NLS of TDP-43) - however a disruption of the KPNB proteins is key to this. Hence poly-GR could disrupt the KPNB proteins, the consequence of which is a mislocalisation of the KPNAs and this KPNA dysfunction could then lead to the subsequent cytoplasmic accumulation of TDP-43. *Drosophila* CAS and *Drosophila* KPNB1 in poly-GR expressing flies could not be tested as the antibodies either did not work or were no longer available. Looking at how *TBPH* localisation changes in salivary gland cells expressing poly-GR following knockdown or overexpression of KPNA and KPNB proteins will help to determine if a dysfunction in the nuclear import receptors does indeed contribute to the cytosolic accumulation of *TBPH*. Poly-GR caused *Drosophila* TDP-43 to accumulate in the cytoplasm without an observable nuclear loss. A nuclear depletion of TDP-43 may take time to develop and hence may not be observable in the early larval time point studied. Indeed Nishimura et al. (2010) found KPNB1 knockdown caused a nuclear depletion and cytoplasmic accumulation of KPNA2, which was accompanied by a concomitant cytoplasmic accumulation of TDP-43 in cells – however, no loss of nuclear TDP-43 was seen in these KPNB1 siRNA treated cells. In another *Drosophila* model that looked at *TBPH* localisation in salivary gland cells expressing G_4C_2 repeats, the protein was found to accumulate in the cytoplasm but was not depleted from the

nucleus (Zhang et al., 2015). Hence nuclear loss may become apparent over time and be a secondary event after cytoplasmic accumulation.

The finding that the *Drosophila* KPNAs are depleted from the nucleus at this early larval stage (before TDP-43 is depleted from the nucleus) suggests that KPNA/B dysfunction may be an early initiating event in C9ALS/FTD pathogenesis. In a review of the literature combined with bioinformatics analysis by Boeynaems et al., (2016b) suggested that disturbances in nucleocytoplasmic transport play a vital initiating role in the development of ALS/FTD. KPNA/B dysfunction is not the only explanation for the cytoplasmic accumulation of *Drosophila* TDP-43 observed following poly-GR expression. As noted the arginine DPRs disrupt in the function of the NPC itself making it more viscous, potentially disrupting import and export through the nuclear pore (Shi et al., 2017). Hence the cytoplasmic accumulation of *TBPH* could result from an inability of the protein to move through the NPC. The KPNA and KPNB proteins may be particularly sensitive to defecits in the NPC due to their constant shuttling through the nuclear pore. Khosravi et al. (2017) found the expression of KPNA3, NUP54 and NUP62, but not RAN or RanGAP, fully restored nuclear localization of the NLS-TDP reporter to control levels further suggesting disturbance in the dynamics of NPC and nuclear import receptors is important to TDP-43 mislocalisation. Interestingly Nishimura et al. (2010) found knockdown of KPNB1 and CAS cause a more severe cytoplasmic accumulation of TDP-43 than a knockdown of Nups suggesting KPNA/B dysfunction is particularly important to TDP-43 mislocalisation. Poorly dynamic SGs induced as a consequence of poly-GR expression (Lee et al., 2016) could also contribute to TDP-43 and KPNA mislocalisation. As discussed in this thesis both TDP-43 and the KPNA/B proteins are SG components (Liu-Yesucevitz et al., 2010; Bentmann et al., 2012; McGurk et al., 2014; Mahboubi et al., 2013). TDP-43 and nuclear import receptors may become trapped in these poorly dynamic SGs in the cytoplasm. Some cellular models have observed poly-GR and poly-PR cause TDP-43 to co-localise with SG markers (Tao et al., 2015; Lee et al., 2016). It will be interesting to look whether KPNA/B proteins are present in the poorly dynamic SGs induced by poly-GR and if they colocalise with TDP-43 in these SGs.

Finally, it was also demonstrated that *Drosophila* TDP-43 and *Drosophila* KPNA2, KPNA4 and RCC1 were sequestered into poly-GA inclusions in the cytoplasm. Poly-GA inclusions were p62 positive; suggestive of defective autophagy or proteosome dysfunction. Poly-GA has previously been associated with nucleocytoplasmic transport dysfunction; Zhang et al. (2016) found nucleocytoplasmic transport proteins are sequestered into poly-GA inclusions in poly-GA expressing mice. Khosravi et al. (2017) found poly-GA cytoplasmic aggregates cause a TDP-NLS reporter to be mislocalised to the cytoplasm; however, the overexpression of KPNA3 and KPNA4, the receptors for the classical NLS of TDP-43, rescued this. Poly-GA has also been previously shown to disrupt UPS function (Zhang et al., 2014; Zhang et al., 2016); poly-GA induced defects in protein degradation may cause this KPNA2 and KPNA4 inclusion formation in the cytoplasm. Poly-GA may sequester KPNAs leading to an exclusion of TDP-43 from the nucleus; making it more vulnerable to be sequestered by cytoplasmic poly-GA. Further poly-GA UPS dysfunction may directly promote the aggregation of TDP-43.

Interestingly both poly-GA and poly-GR disrupt the localisation of nuclear import proteins yet only poly-GR shows severe toxicity. In poly-GA expressing salivary glands nuclear import proteins are sequestered into inclusions in the cytoplasm but are not depleted from nucleus – whereas in poly-GR expressing salivary glands there is a nuclear depletion of said nuclear import proteins. This suggests a loss of nuclear function and/or an inability of these proteins to shuttle between the nucleus and cytoplasm could be the potential cause of the severe toxicity seen following poly-GR expression. For poly-GA, although *Drosophila* KPNA2 and KPNA4 are present in poly-GA aggregates, non-sequestered functional KPNA2 and KPNA may still be able to shuttle through the nuclear pore. Indeed, both KPNAs are still present in the nucleus in poly-GA expressing salivary glands indicating there is still enough of the non-sequestered protein to circumvent the detrimental effects of KPNA sequestration into poly-GA cytoplasmic inclusions. It is possible over time toxicity might develop as the number of aggregates increase and more of these import proteins become trapped in poly-GA aggregates.

The question therefore is why poly-GR causes a depletion of *Drosophila* KPNA2 and KPNA4 from the nucleus and not poly-GA. One explanation could be a direct

interaction between soluble poly-GR and these karyopherins in cytoplasm which prevents then from completely shuttling into the nucleus. KPNA2 was present in a poly-GR/poly-PR protemoics screen performed by Lee et al., (2016), furthermore pulldown demonstrated there is a a direct interaction between these DPRs and KPNA2. Poly-GR may have a stronger interaction with these nuclear import proteins in the cytoplasm than poly-GA. Nuclear import proteins are SG components (Fujimura et al., 2010; Markmiller et al., 2018) and poly-GR causes SGs to form which are unable to disassemble (Lee et al., 2016). Future work should therefore investigate if KPNA2 and KPNA4 are present in cytoplasmic SGs in poly-GR expressing *Drosophila* tissue. Another potential explanation for this karyopherin nuclear depletion could be poly-GR induced disrupions of nuclear pore function. No morphological abnormalities of the NPC were observed in this study following poly-GR expression, however this does not mean the function of the pore itself is not disturbed. Like membraneless organelles, such as SGs, nuclear pore proteins are known to phase transition from soluble aqueous to hydrogel-like states (Schmidt and Gorlich, 2016). The FG-Nups that undergo these phase seperations into a hydrogel are responsible for the selectivity of the NPC by providing a permeability barrier. The arginine DPRs drisrupt phase separations of proteins with intrinsically disordered regions and the FG domains of these FG-Nups share features of these intrinsically disordered regions such as low hydrophobicity and high net charge (Schmidt and Gorlich, 2016). Hence poly-GR may alter biophysical properties of the hydrogels disrupting KPNA2 and KPNA4 import into the nucleus. Indeed poly-PR, binds to the FG-repeats in the central channel of the pore making the NPC less permeable (Shi et al., 2017).

Furthermore, the different mechanisms by which poly-GR and poly-GA could disrupt nucleocytoplasmic transport (perturbed SG dynamics and NPC deficits for poly-GR and UPS dysfunction for poly-GA) may explain why a cytoplasmic accumulation of TBPH is seen from poly-GR expression whereas poly-GA causes TBPH inclusions in the cytosol. Interestingly when poly-GA and TDP-43 inclusions co-aggregate in the human brain; poly-GA is found at the centre of these inclusions suggesting it seeds TDP-43 accumulation (Mori et al., 2013a; Mackenzie et al., 2013). Lee et al. (2017) found poly-GA is also found in the centre of poly-GP/poly-GA co-aggregates, further indicating poly-GA sequesters other proteins into aggregates. One caveat is that

poly-GA aggregates and TDP-43 pathology do not correlate in the human post-mortem brain (Mackenzie et al., 2015); however, poly-GA has been shown to be transmissible (Chang et al., 2016; Westergard et al. 2016) which may explain the lack of spatial correlation between poly-GA inclusions and TDP-43 pathology (Edbauer and Haass, 2016). A proposed model for poly-GR mediated cytoplasmic accumulation of KPNAs and TDP-43 is presented in figure 9.1.

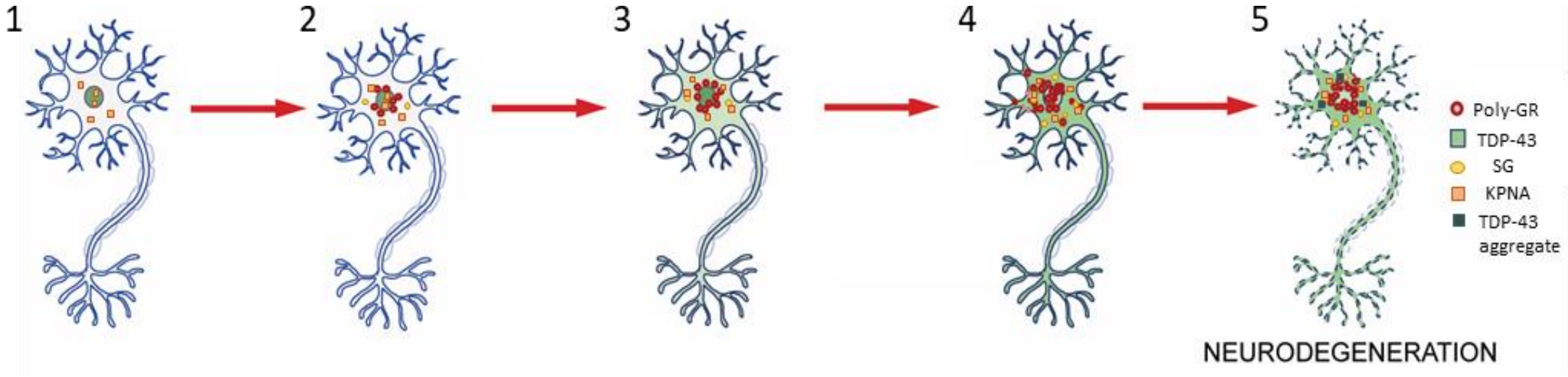

Figure 9.1 | Proposed model for poly-GR mediated cytoplasmic accumulation of KPNAs and TDP-43. (1) In a healthy neuron TDP-43 is predominately localised in the nucleus, the karyopherins (KPNAs) shuttle between the cytoplasm and the nucleus. The KPNAs are involved in the classical nuclear import pathway, they directly bind to the NLS of TDP-43 and shuttle the protein between the cytoplasm and nucleus **(2)** Poly-GR causes the KPNAs to be depleted from the nucleus and accumulate in the cytoplasm. There are several mechanisms by which poly-GR may induce a mislocalisation of these kayropherins. Cytoplasmic poly-GR may directly bind to these nuclear import proteins in the cytoplasm preventing them entering the nucleus. Poly-GR may also disrupt the function of the nuclear pore complex, possibly by altering the phase separations of the FG-nucleoporin hydrogels, thereby perturbing nuclear import of the KPNAs leading to their accumulation in the cytoplasm. Poly-GR also causes the formation of stress granules (SGs) which are unable to disassemble, as KPNAs are known components of SGs they could become trapped in these cytoplasmic membraneless organelles preventing their entry to and from the nucleus. **(3)** The depletion of the KPNAs into the cytoplasm disrupts the normal nuclear import of TDP-43 causing it to accumulate in the cytoplasm. **(4)** Cytoplasmic TDP-43 causes an increase in RAN translation leading to higher amounts of poly-GR of in the cytoplasm further exasperating poly-GR mediated KPNA nuclear depletion. This will lead to a further cytoplasmic accumulation of TDP-43 and eventually a nuclear depletion as well. **(5)** As a known SG component, soluble cytoplasmic TDP-43 will also associate with perturbed SGs becoming insoluble eventually leading to TDP-43 aggregate formation. A combination of cytoplasmic TDP-43 gain of function and toxic nuclear loss of TDP-43 function leads to neurodegeneration.

9.3 KPNA4 is mislocalised in both C9FTD and sporadic FTD-TDP frontal cortex

The present study found KPNA4 localisation (the human homologue of *Drosophila importin-α3*) was disrupted in sporadic FTD-TDP and C9FTD human frontal cortex tissue. KPNA4 appeared downregulated in disease compared to controls in frontal cortex tissue, although the N number was too low for statistical significance. Furthermore, KPNA4 was present in the urea insoluble fraction for sporadic FTD-TDP and C9FTD suggesting it may be forming insoluble inclusions. Finally, KPNA4 localisation was perturbed in both sporadic FTD-TDP and C9FTD frontal cortex compared to controls. In controls the protein was distributed across the cytoplasm and nucleus with an enrichment in the cytoplasm. KPNA4 in sporadic FTD-TDP and C9FTD frontal cortex neurons was either depleted from the nucleus or accumulated within the nucleus. Nishimura et al. (2010) found CAS was downregulated in sporadic FTD-TDP frontal cortices; CAS binds to the KPNAs in the nucleus recycling them into the cytoplasm, hence the nuclear accumulation of KPNA4 may be caused by an inhibition of CAS mediated KPNA4 nuclear export. The nuclear depletion could result from KPNB1 dysfunction which binds KPNA4 in the cytosol. The disruptions in the function the NPC and SG dynamics caused by the arginine DPRs may provide an explanation for the KPNA4 mislocalisation and TDP-43 aggregation seen in C9ALS/FTD frontal cortex as discussed. Although additional post-mortem work is needed to investigate whether arginine DPR pathology correlates with the KPNA4 mislocalisation observed. However, how might KPNA4 and TDP-43 become mislocalised in sporadic FTD-TDP where there is no poly-GR to disrupt the NPC or SG dynamics?

There has been an increasing recognition over a number years at nucleocytoplasmic transport defects have a role in ALS and FTD cases, including cases not caused by the *C9ORF72* repeat expansion (Kim and Taylor, 2017). For example, half of the mutations in FUS associated with ALS and FTD are located in the C-terminal non-classical proline-tyrosine nuclear localisation signal (PY-NLS) (Dormann and Haass, 2011; Kim and Taylor, 2017). Further the extent to which these mutations disrupt the nuclear localisation of FUS correlates with both reduced onset and enhanced progression of disease (Bosco et al., 2010; Chiò et al., 2009; DeJesus-Hernandez et

al., 2010; Dormann et al., 2010). A missense mutation in the NLS of hnRNPA1 is a cause of familial ALS, this mutation leads to a nuclear depletion and cytosolic accumulation of hnRNPA1 (Liu et al., 2016b). Mutations in GLE1, a component of the NPC, have been identified in two motor neuron diseases closely related to ALS - lethal congenital contracture syndrome 1 and lethal arthrogryposis with anterior horn cell disease (Nousiainen et al., 2008). Further a rare GLE1 loss of function mutation has also been recently identified in French and French-Canadian ALS patients (Kaneb et al., 2015). As just noted the protein levels and localisation of members of the classic importin α/ β pathway such as KNPA2 and CAS are altered in ALS and FTD tissue (Nishimura et al., 2010). Trn1 inclusions have also been identified in FTD-FUS cases and adult onset ALS without FUS mutations (Neumann et al., 2012; Takeuchi et al., 2013).

Evidence suggests cerebral ageing has a detrimental impact on the NPC (Kim and Taylor, 2017) (figure 9.2). Nucleocytoplasmic transport is thought to be one the cellular processes affected most by age-related decline, specifically the integrity of the NPC (Kim and Taylor, 2017). NPC proteins in neurons are extremely long lived and remain in the NPC for the whole lifespan of an organism (Savas et al., 2012; Toyarna et al., 2013). Kim and Taylor, (2017) propose this is because in dividing cells the NPC disassembles during mitosis, reassembling with newly synthesized proteins; however, as neurons are post mitotic the complete disassembly and reassembly is not practical, rather NPC proteins are more likely to be slowly replaced as a neuron ages for newly translated NPC proteins. This however makes the NPC specifically vulnerable to damage from harmful stressors (Kim and Taylor et al, 2017).

For example, D'Angelo et al. (2009) found age-dependent deterioration of NPCs, via oxidative mediated damage; which led to an increase in the permeability of the nuclear pore causing changes in the subcellular concentrations of cytoplasmic proteins. Interestingly age-dependent changes in the protein concentrations of nuclear transport receptors KPNA2 and CAS have also been reported in fibroblasts (Pujol et al*.,* 2002). Mertens et al. (2015) generated induced neurons from human donors from a broad range of ages, finding aged induced neurons had specific transcriptional profiles including decreases in RanBP17 levels. Furthermore, induced neurons from aged donors had impaired nucleocytoplasmic transport. Reducing

RanBP17 levels in induced neurons from young donors impaired nucleocytoplasmic transport leading to similar transcriptional changes seen in induced neurons from aged donors, further solidifying the importance of nucleocytoplasmic transport in human ageing. Therefore, one hypothesis for sporadic forms of disease is that specific genetic, environmental and molecular stressors make aged neuronal populations which are highly susceptible to stress unable mount an adequate stress response. One of the first cellular processes to be disrupted is nucleocytoplasmic transport due to age-related impairments in the NPC, leading to a disruption in the localisation of proteins that shuttle between the nucleus and cytoplasm, including KPNA and KNPB proteins and also TDP-43.

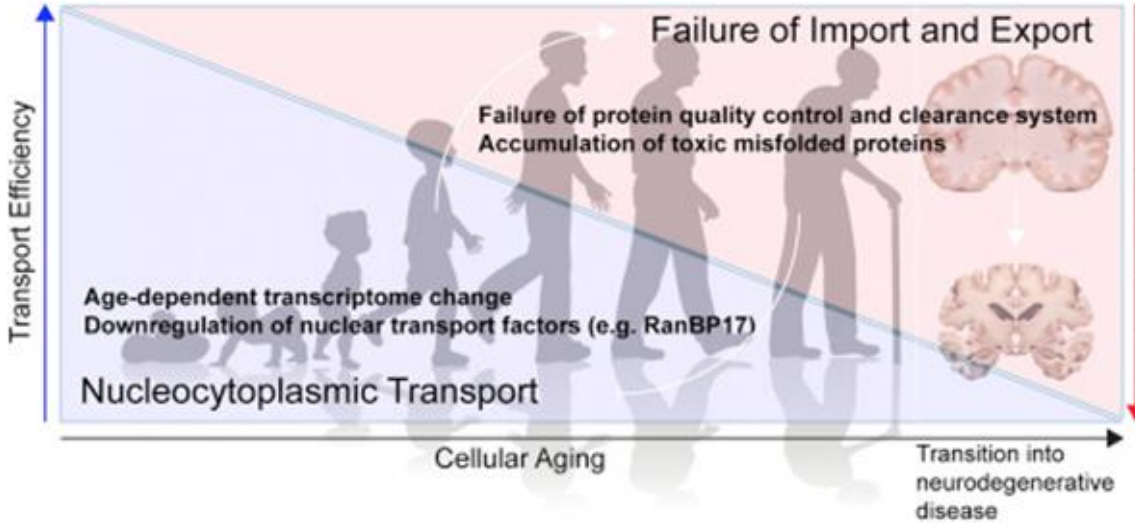

Figure 9.2 | Cellular Aging, disease, and nucleocytoplasmic transport. Disrupted nucleocytoplasmic transport has been reported as both a consequence of age and in neurodegenerative diseases such as ALS. An age-dependent reduction in levels of proteins involved the transport of molecules across the nuclear pore complex, such as the karyopherins and RanBP17 leads to perturbed nucleocytoplasmic transport resulting in the aggregation of proteins in the cytosol which further impairs nuclear import and export via the sequestration of nucleocytoplasmic transport factors. From Kim and Taylor, (2017)

A recent landmark paper by Woerner et al. (2016) demonstrating the pathological aggregation of β-sheet proteins in cytoplasm impairs nuclear import and export may explain the disruption of KPNA4 in sporadic FTD-TDP. Woerner et al (2016) targeted artificial β-sheet proteins to either the nucleus or cytoplasm using NES or NLS sequences respectively. Cytoplasmic β protein aggregates, but not nuclear β protein

aggregates were highly toxic to cells. Woerner et al. (2016) investigated why there were such differences in toxicity between β proteins located in different subcellular compartments by looking at their structural properties. Using NIAD-4, a fluorescent sensor that binds to amyloid structures; it was found that nuclear β protein aggregates only weakly bind the amyloid sensor NIAD-4 compared to their cytoplasmic counterparts suggesting significant structural differences between cytoplasmic and nuclear aggregates.

To uncover mechanistically why cytoplasmic, but not nuclear, β proteins were so toxic Woerner et al. (2016) probed the localisation behaviour of the β proteins by staining for NPC proteins. Using antibodies specific for FG-repeat Nups (e.g. Nup, 62, Nup152, Nup90) these FG-repeat Nups were found to be recruited to cytoplasmic – but not nuclear β protein aggregates. Furthermore, cytoplasmic β protein aggregates caused irregularities in the shape of the nuclear envelope and nucleus itself. None of this was observed for nuclear β protein aggregates. Strikingly Woerner et al. (2016) also observed human disease proteins forming predominantly cytoplasmic aggregates, Htt96Q, TDP-43 and Parkin – also perturbed the NPC, with aggregates of FG-repeat Nups seen in the cytoplasm – although they did not frequently co-localise with human transgenic proteins with the artificial cytoplasmic β protein aggregates. This data therefore suggests cytoplasmic β protein aggregates disrupt nuclear cytoplasmic transport.

To explore this further Woerner et al. (2016) fused an EGFP to both an NLS and NES, meaning EGFP is continuously shuttled between the nucleus and cytoplasm. The fluorescence of this shuttle GFP (S-GFP) was quantified in the nucleus and cytoplasm respectively to investigate defective nuclear cytoplasmic transport as a result of β protein aggregates. At equilibrium S-GFP is predominantly cytoplasmic indicating nuclear export is faster than nuclear import. Nuclear β protein aggregates had no effect on the distribution of S-GFP; however cytoplasmic β protein aggregates caused a significant increase in nuclear S-GFP suggesting defective nuclear export. This was also observed for cytoplasmic HTT96Q and cytoplasmic ParkinΔC aggregates co-expressing with S-GFP. When the exportin inhibitor leptomycin B (LMB) was applied to cells S-GFP translocated to the nucleus within 15 minutes indicating quick nuclear import. This nuclear import was inhibited by cytoplasmic β protein aggregates and Htt96Q, but not by nuclear β protein

aggregates. Hence cytoplasmic β protein aggregates disrupt both nuclear import and export.

To explore these defects further Woerner et al. (2016) investigated the transport of NF-κB. NF-κB is a protein complex that functions as a transcription factor transported into the nucleus by nuclear import receptors – specifically members of the KPNA α2 family – KPNA3 and KPNA4. In control cells and cells with nuclear β protein aggregates NF-κB was transported into the nucleus. However, in cells expressing cytoplasmic β protein aggregates and Htt96Q, TDP-43, and ParkinΔC the nuclear import of NF-κB was defective – the protein remained cytoplasmic, it did not translocate into the nucleus and was sometimes recruited into cytoplasmic β protein aggregates. This deficit seemed specific to the nuclear import of NF-κB as the upstream NF-κB signalling pathway was not affected. To investigate defective nuclear import Woerner et al. (2016) looked at the main import carrier for NF-κB – KPNA4. In control cells KPNA4 was distributed in both the cytoplasm and nucleus with a stronger cytoplasmic stain. Cytoplasmic β protein aggregates sequestered KPNA4 into inclusions; Htt96Q or ParkinΔC aggregates also caused KPNA4 aggregate formation but these aggregates did not co-localise with human transgene protein aggregates. Similar results were seen for KPNA2. Both KPNAs were unaffected by nuclear β protein aggregates. The authors however did not look at TDP-43 and the KPNAs in their cellular model.

The Woerner et al. (2016) paper shows a clear dysfunction and co-aggregation of KPNAs as a consequence of cytoplasmic protein aggregates. The data from the present study shows in human post-mortem tissue that KPNA4 is potentially sequestered into TDP-43 aggregates. Woerner et al. (2016) found expression of aggregate prone mutant TDP-43 causes a cytoplasmic accumulation of nuclear THOC2 – a protein involved in mRNA export. The authors propose transport factors may specifically be affected by cytoplasmic aggregates as they often contain disordered and low complexity domains causing aberrant interactions with the interactive surfaces of cytosolic aggregates.

The Woerner et al. (2016) study therefore demonstrates that cytoplasmic aggregates of β sheet forming human disease proteins impair nucleocytoplasmic transport. This study has a number of important implications with regards to the development of

TDP-43 proteinopathy and its relationship with defective nucleocytoplasmic transport in sporadic forms of ALS and FTD. One interpretation is that some sort of feed-back loop exists in which defective nucleocytoplasmic transport, caused by age-related deficits in the NPC, leads to TDP-43 accumulation in the cytoplasm. The subsequent aggregation of TDP-43 further disrupts nucleocytoplasmic transport creating a vicious cycle. Hence KPNA4 mislocalisation in sporadic FTD-TDP cases could be a result of TDP-43 mediated deficits in nucleocytoplasmic transport. A nuclear depletion of KPNA4 is in sporadic FTD-TDP frontal cortex was not however always accompanied by a TDP-43 aggregate. Hence aggregates of TDP-43 cannot per-say be the sole cause of the KPNA4 mislocalisation observed. TDP-43 not only binds nucleocytoplasmic transport proteins such as KPNA4, but it may also regulate the splicing and transcription of nucleocytoplasmic transport proteins leading to downstream defective nuclear import (Dormann and Haass, 2011). Hence a loss of TDP-43 function via cytoplasmic accumulation (without aggregation) and/or nuclear depletion could also disrupt nucleocytoplasmic transport. This raises the question as to whether initial modest TDP-43 mislocalisation can impair nuclear transport on its own and form a positive feedback loop, causing more TDP-43 to accumulate and aggregate in the cytoplasm.

Several studies have demonstrated that TDP-43 levels directly impact nuclear transport; indicating some kind of feed-back regulation mechanism (Mihevc et al., 2017). Proteomic analysis has revealed TDP-43 knockdown affects the levels of several proteins involved nuclear import and export including Nup153, Nup205, Nup160, RAE1, RANBP1, CRM1, IPO5, IPO7, calreticulin and POM121C (Štalekar et al., 2015). A loss of TDP-43 has been shown to cause a reduction in RAN levels and non-functional RAN mutants lead to a decrease in the amount of nuclear TDP-43 (Ward et al., 2014). TDP-43 regulates RAN expression via binding to the 3'UTR of its mRNA, hence a loss of nuclear TDP-43 could lead to a downregulation of RAN which would further cause TDP-43 nuclear loss and its subsequent cytoplasmic accumulation (Mihevc et al., 2017). Hence it will be interesting to investigate the localisation and function of nucleocytoplasmic transport in model systems in which TDP-43 is mislocalised without the formation of aggregates. It would also be intriguing to stain for WT TDP-43 in the human FTD frontal cortex to see if the KPNA4 mislocalisation correlated with a cytoplasmic

accumulation/nuclear depletion of TDP-43. Another interesting point to note is that poly-GA forms a parallel β-sheet structure with similar structural properties to the amyloid-beta protein in Alzheimer's disease (Chang et al., 2016; Freibaum and Taylor, 2017). Khosravi et al. (2017) found cytoplasmic, but nuclear poly-GA aggregates cause a mislocalisation of TDP-43. Hence the poly-GA induced aggregation of *Drosophila* KPNA2, KPNA4 and RCC1 could be mediated by similar mechanisms to what is reported for other cytoplasmic β proteins by Woerner et al. (2016).

The Woerner et al. (2016) study also suggests the accumulation and aggregation of TDP-43 in the cytoplasm could precede defective nucleocytoplasmic transport. Therefore, what could be the initial trigger of TDP-43 aggregation? Dormann and Haass, (2011) proposed nucleocytoplasmic transport could be a secondary phenomenon resulting from another primary insult. Another cellular process known also to be affected by ageing is the proteostasis machinery (Kim and Taylor, 2017). An age-related decline in proteostasis involves defects in the stabilization of correctly folded proteins and also protein clearance systems (Saeza and Vilchez, 2014). For example, the induction of chaperones, involved in cellular processes that maintain proper protein folding in response to stress, such as the heat shock response and unfolded protein response, is impaired in ageing (Calderwood et al., 2009). Further when chaperones are overexpressed *in vivo* life span is increased (Morley and Morimoto, 2004; Morrow et al., 2004). In mice mutant strains with long life-spans show high levels of heat shock chaperones while mice lacking members of the heat shock protein family show increased ageing (Min et al., 2008; Swindell et al., 2009; Saeza and Vilchez, 2014). Further both defects in the ubiquitin proteasome system (UPS) and autophagy are associated with age (Morimoto and Cuervo, 2014; Saeza and Vilchez, 2014). Ageing has been shown to cause several impairments in proteasome function including a reduction in the expression of proteasome subunits and the disassembly of proteasomes (Vernace et al., 2007; Saeza and Vilchez, 2014). In yeast age-related UPS dysfunction causes protein aggregation (Andersson et al., 2013). Reduced proteasome function as a consequence of ageing has been observed in the human cerebral cortex, hippocampus and spinal cord suggesting the activity of the proteasome declines with age (Saeza and Vilchez, 2014). Markers of autophagy are also reduced in the aged human brain (Saeza and Vilchez, 2014) and

using rapamycin to induce autophagy delays ageing *in vivo* and promotes lifespan (Harrison et al., 2009; Bjedov et al., 2010; Wilkinson et al., 2012).

Impaired proteasome function is associated with sporadic ALS (Kabashi et al., 2012); autophagy is also upregulated and activated sporadic ALS spinal cord motor neurons (Sasaki, 2011). There are links to defective proteostasis in familial forms ALS/FTD associated with TDP-43 proteinopathy such as mutations in *GRN*, *VCP, SQSTM1, UBQLN2, OPTN, and TBK1*. All of these genes are involved in autophagy and lysosomal pathways (Hardy and Rogaeva, 2014). It is possible that age related deficits proteostasis may be a bigger contributor to the development of TDP-43 aggregates in sporadic cases than disrupted SG dynamics which may be moe important in C9ALS/FTD. Indeed McGurk et al. (2014) investigated the SG marker PABP-1 in human spinal cord motor neurons finding the colocalization of PABP-1 and TDP-43 was twice as frequent in ALS patients with the *C9ORF72* expansion compared to ALS with no mutation.

9.4 Outlook

A major question in ALS/FTD is how RNA binding proteins, in particular TDP-43, become depleted from the nucleus and sequestered into the cytoplasm (Dormann and Haass, 2011). A multi-hit hypothesis for TDP-43 proteinopahy has been proposed by Dormann and Haass, (2011) in which the development of TDP-43 inclusions involves defective nuclear import as an initiating factor followed by SG mediated aggregation. For *C9ORF72* cases the mechanisms underpinning the pathway toward TDP-43 proteinopathy are becoming clearer. The arginine DPRs provide both of these proposed hits; their binding to the LCD of proteins and subsequent disruption of phase transitions disrupts both the NPC (which will disrupt nuclear import) and SG dynamics, leading to SGs which do not disassemble, argued to be the basis of TDP-43 inclusion formation (Becker et al., 2017). Poly-GA may also have a role through defective proteostasis and the sequestration of nucleocytoplasmic transport proteins into its inclusions. The C9ORF72 protein could also contribute due its roles in autophagy (Webster et al., 2016a; Sellier et al., 2016) - the mechanism by which the DPRs are cleared by (Cristofani et al., 2017), and also nucleocytoplasmic transport (Xiao et al., 2015a) and SG formation (Maharjan et al., 2017). The G_4C_2 RNA has also been shown to bind to and sequester proteins

involved in nucleocytoplasmic transport (Zhang et al., 2015). Further the repeat RNA has recently been shown to also undergo phase transitions (Jain and Vale, 2017) and it will be interesting to see what role this has in toxicity.

There are some caveats to DPR mediated disruptions in the nucleocytoplasmic transport-SG axis as the cause of TDP-43 pathology in C9ALS/FTD. How can rare C9ALS/FTD cases which have DPR pathology, but no TDP-43 pathology be explained? Also, if poly-GR and poly-PR perturb phase transitions of RNA binding proteins why is the pathology always so specific to TDP-43. Why is FUS or hnRNPA1 pathology not seen in the C9ALS/FTD brain? Both proteins undergo phase transitions and are SG components like TDP-43 (Patel et al., 2015; Mollieux et al., 2015). A further understanding of the structural and functional differences between RNA binding such as FUS and TDP-43 may help answer this question. Also, what is behind the development of TDP-43 pathology in sporadic cases? Recent studies suggest TDP-43 pathology itself perturbs nucleocytoplasmic transport (Woerner et al., 2016) suggesting cellular processes other than disrupted nuclear import could cause TDP-43 aggregation; which age-related impairments in proteostasis being one potential mechanism. A better understanding of environmental and genetic risk factors that make some more susceptible to developing TDP-43 pathology and how these factors disrupt this nuclear import in combination with ageing will hopefully shed light on this open question.

Chapter 10: Conclusion

This thesis reports a *Drosophila* model of C9ALS/FTD characterised by motor impairment that varied in age of onset and severity. The characteristics of these motor deficits were dependent on the levels and identity of dipeptide repeat proteins (DPRs) produced, and not the G_4C_2 RNA, with severe, early onset phenotypes correlating with poly-GR. Expression of poly-GR leads to a cytoplasmic accumulation of *Drosophila* TDP-43, indicating poly-GR is the key stressor in the disease cascade, accumulating early during the disease progression and subsequently causing TDP-43 to be mislocalised to the cytoplasm providing a direct link between two major pathological hallmarks of C9 ALS/FTD. Further poly-GA also recruited *Drosophila* TDP-43 into inclusions. Moreover, not only does cytoplasmic TDP-43 enhance G_4C_2 toxicity and is sufficient itself for toxicity but show DPR levels are also enhanced by the targeted increase in cytoplasmic TDP-43. Furthermore, both poly-GR and poly-GA disrupt the localisation of the nuclear import protein *Importin-α3* and *Pendulin* in flies suggesting defective nuclear import. The human homologue *Importin-α3*, KPNA4 was also mislocalised in sporadic FTD-TDP and C9FTD frontal cortex. Thus, these findings suggest the DPRs cause TDP-43 proteinopathy underlying C9ALS/FTD.

References

Aguzzi, A., Altmeyer, M., 2016. Phase Separation: Linking Cellular Compartmentalization to Disease. Trends Cell Biol.

Akimoto, C., Volk, A.E., van Blitterswijk, M., Van den Broeck, M., Leblond, C.S., Lumbroso, S., Camu, W., Neitzel, B., Onodera, O., van Rheenen, W., Pinto, S., Weber, M., Smith, B., Proven, M., Talbot, K., Keagle, P., Chesi, A., Ratti, A., van der Zee, J., Alstermark, H., Birve, A., Calini, D., Nordin, A., Tradowsky, D.C., Just, W., Daoud, H., Angerbauer, S., DeJesus-Hernandez, M., Konno, T., Lloyd-Jani, A., de Carvalho, M., Mouzat, K., Landers, J.E., Veldink, J.H., Silani, V., Gitler, A.D., Shaw, C.E., Rouleau, G.A., van den Berg, L.H., Van Broeckhoven, C., Rademakers, R., Andersen, P.M., Kubisch, C., 2014. A blinded international study on the reliability of genetic testing for GGGGCC-repeat expansions in C9orf72 reveals marked differences in results among 14 laboratories. J. Med. Genet. 51, 419–424.

Akizuki, M., Yamashita, H., Uemura, K., Maruyama, H., Kawakami, H., Ito, H., Takahashi, R., 2013. Optineurin suppression causes neuronal cell death via NF-kB pathway. J. Neurochem. 126, 699–704.

Alfieri, J.A., Pino, N.S., Igaz, L.M., 2014. Reversible behavioral phenotypes in a conditional mouse model of TDP-43 proteinopathies. J. Neurosci. 34, 15244–59.

Almeida, S., Gascon, E., Tran, H., Chou, H.J., Gendron, T.F., Degroot, S., Tapper, A.R., Sellier, C., Charlet-Berguerand, N., Karydas, A., Seeley, W.W., Boxer, A.L., Petrucelli, L., Miller, B.L., Gao, F.B., 2013. Modeling key pathological features of frontotemporal dementia with C9ORF72 repeat expansion in iPSC-derived human neurons. Acta Neuropathol. 126, 385–399.

Al-Sarraj, S., King, A., Troakes, C., Smith, B., Maekawa, S., Bodi, I., Rogelj, B., Al-Chalabi, A., Hortobágyi, T., Shaw, C.E., 2011. P62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTLD and MND/ALS. Acta Neuropathol. 122, 691–702.

Amick, J., Ferguson, S.M., 2017. C9orf72: At the intersection of lysosome cell biology and neurodegenerative disease. Traffic.

Amick, J., Roczniak-Ferguson, A., Ferguson, S.M., 2016. C9orf72 binds SMCR8, localizes to lysosomes, and regulates mTORC1 signaling. Mol. Biol. Cell 27, 3040–3051.

Amlie-Wolf, A., Ryvkin, P., Tong, R., Dragomir, I., Suh, E., Xu, Y., Van Deerlin, V.M., Gregory, B.D., Kwong, L.K., Trojanowski, J.Q., Lee, V.M.Y., Wang, L.S., Lee, E.B., 2015. Transcriptomic changes due to cytoplasmic TDP-43 expression reveal dysregulation of histone transcripts and nuclear chromatin. PLoS One 10, e0141836.

Andersson, V., Hanzén, S., Liu, B., Molin, M., Nyström, T., 2013. Enhancing protein disaggregation restores proteasome activity in aged cells. Aging (Albany. NY). 5, 802–812.

Aoki, Y., Manzano, R., Lee, Y., Dafinca, R., Aoki, M., Douglas, A.G.L., Varela, M.A., Sathyaprakash, C., Scaber, J., Barbagallo, P., Vader, P., Mäger, I., Ezzat, K., Turner, M.R., Ito, N., Gasco, S., Ohbayashi, N., El Andaloussi, S., Takeda, S., Fukuda, M., Talbot, K., Wood, M.J.A., 2017. C9orf72 and RAB7L1 regulate vesicle trafficking in amyotrophic lateral sclerosis and frontotemporal dementia. Brain 140, 887–897.

Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., Mann, D., Tsuchiya, K., Yoshida, M., Hashizume, Y., Oda, T., 2006. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem. Biophys. Res. Commun. 351, 602–611.

Arnold, E.S., Ling, S.-C., Huelga, S.C., Lagier-Tourenne, C., Polymenidou, M., Ditsworth, D., Kordasiewicz, H.B., McAlonis-Downes, M., Platoshyn, O., Parone, P.A., Da Cruz, S., Clutario, K.M., Swing, D., Tessarollo, L., Marsala, M., Shaw, C.E., Yeo, G.W., Cleveland, D.W., 2013. ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43. Proc. Natl. Acad. Sci. 110, E736–E745.

Arnold, W.D., Kassar, D., Kissel, J.T., 2015. Spinal muscular atrophy: diagnosis and management in a new therapeutic era. Muscle Nerve 51, 157–67.

Ash, P.E.A., Bieniek, K.F., Gendron, T.F., Caulfield, T., Lin, W.L., DeJesus-Hernandez, M., Van Blitterswijk, M.M., Jansen-West, K., Paul, J.W., Rademakers, R., Boylan, K.B., Dickson, D.W., Petrucelli, L., 2013. Unconventional Translation of C9ORF72 GGGGCC Expansion Generates Insoluble Polypeptides Specific to c9FTD/ALS. Neuron 77, 639–646.

Atanasio, A., Decman, V., White, D., Ramos, M., Ikiz, B., Lee, H.-C., Siao, C.-J., Brydges, S., LaRosa, E., Bai, Y., Fury, W., Burfeind, P., Zamfirova, R., Warshaw, G., Orengo, J., Oyejide, A., Fralish, M., Auerbach, W., Poueymirou, W., Freudenberg, J., Gong, G., Zambrowicz, B., Valenzuela, D., Yancopoulos, G., Murphy, A., Thurston, G., Lai, K.-M.V., 2016. C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production and glomerulonephropathy in mice. Sci. Rep. 6, 23204.

Atkinson, R.A.K., Fernandez-Martos, C.M., Atkin, J.D., Vickers, J.C., King, A.E., 2015. C9ORF72 expression and cellular localization over mouse development. Acta Neuropathol. Commun. 3, 59.

Aulas, A., Vande Velde, C., 2015. Alterations in stress granule dynamics driven by TDP-43 and FUS: a link to pathological inclusions in ALS? Front. Cell. Neurosci. 9, 423.

Ayala, Y.M., Zago, P., D'Ambrogio, A., Xu, Y.-F., Petrucelli, L., Buratti, E., Baralle, F.E., 2008. Structural determinants of the cellular localization and shuttling of TDP-43. J. Cell Sci. 121, 3778–3785.

Baborie, A., Griffiths, T.D., Jaros, E., Perry, R., McKeith, I.G., Burn, D.J., Masuda-Suzukake, M., Hasegawa, M., Rollinson, S., Pickering-Brown, S., Robinson, A.C., Davidson, Y.S., Mann, D.M.A., 2015. Accumulation of dipeptide repeat proteins predates that of TDP-43 in frontotemporal lobar degeneration associated with hexanucleotide repeat expansions in C9ORF72 gene. Neuropathol. Appl. Neurobiol. 41, 601–612.

Bañez-Coronel, M., Ayhan, F., Tarabochia, A.D., Zu, T., Perez, B.A., Tusi, S.K., Pletnikova, O., Borchelt, D.R., Ross, C.A., Margolis, R.L., Yachnis, A.T., Troncoso, J.C., Ranum, L.P.W., 2015. RAN Translation in Huntington Disease. Neuron 88, 667–677.

Barker, H. V., Niblock, M., Lee, Y.-B., Shaw, C.E., Gallo, J.-M., 2017. RNA Misprocessing in C9orf72-Linked Neurodegeneration. Front. Cell. Neurosci. 11, 195.

Barmada, S.J., Skibinski, G., Korb, E., Rao, E.J., Wu, J.Y., Finkbeiner, S., 2010. Cytoplasmic Mislocalization of TDP-43 Is Toxic to Neurons and Enhanced by a Mutation Associated with Familial Amyotrophic Lateral Sclerosis. J. Neurosci. 30, 639–649.

Bauer, P.O., 2016. Methylation of C9orf72 expansion reduces RNA foci formation and dipeptide-repeat proteins expression in cells. Neurosci. Lett. 612, 204–209.

Beaudoin, J.D., Perreault, J.P., 2013. Exploring mRNA 3′-UTR G-quadruplexes: Evidence of roles in both alternative polyadenylation and mRNA shortening. Nucleic Acids Res. 41, 5898–5911.

Beck, J., Poulter, M., Hensman, D., Rohrer, J.D., Mahoney, C.J., Adamson, G., Campbell, T., Uphill, J., Borg, A., Fratta, P., Orrell, R.W., Malaspina, A., Rowe, J., Brown, J., Hodges, J., Sidle, K., Polke, J.M., Houlden, H., Schott, J.M., Fox, N.C., Rossor, M.N., Tabrizi, S.J., Isaacs, A.M., Hardy, J., Warren, J.D., Collinge, J., Mead, S., 2013. Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. Am. J. Hum. Genet. 92, 345–353.

Becker, L.A., Huang, B., Bieri, G., Ma, R., Knowles, D.A., Jafar-Nejad, P., Messing, J., Kim, H.J., Soriano, A., Auburger, G., Pulst, S.M., Taylor, J.P., Rigo, F., Gitler, A.D., 2017. Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice. Nature 544, 367–371.

Beer, A.M., Cooper-Knock, J., Higginbottom, A., Highley, J.R., Wharton, S.B., Ince, P.G., Milano, A., Jones, A. a., Al-Chalabi, A., Kirby, J., Shaw, P.J., 2015. Intermediate length C9orf72 expansion in an ALS patient without classical C9orf72 neuropathology. Amyotroph. Lateral Scler. Front. Degener. 16, 249–251.

Belzil, V. V., Bauer, P.O., Prudencio, M., Gendron, T.F., Stetler, C.T., Yan, I.K., Pregent, L., Daughrity, L., Baker, M.C., Rademakers, R., Boylan, K., Patel, T.C., Dickson, D.W., Petrucelli, L., 2013. Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. Acta Neuropathol. 126, 895–905.

Belzil, V. V., Katzman, R.B., Petrucelli, L., 2016. ALS and FTD: an epigenetic perspective. Acta Neuropathol.

Bennett, W., Fox, R.S.T., 2015. Problems at the nuclear pore Expansion. Nature 525, 6–7.

Bentmann, E., Haass, C., Dormann, D., 2013. Stress granules in neurodegeneration - Lessons learnt from TAR DNA binding protein of 43 kDa and fused in sarcoma. FEBS J.

Bentmann, E., Neumann, M., Tahirovic, S., Rodde, R., Dormann, D., Haass, C., 2012. Requirements for stress granule recruitment of fused in sarcoma (FUS) and TAR DNA-binding protein of 43 kDa (TDP-43). J. Biol. Chem. 287, 23079–23094.

Bieniek, K.F., van Blitterswijk, M., Baker, M.C., Petrucelli, L., Rademakers, R., Dickson, D.W., 2014. Expanded C9ORF72 Hexanucleotide Repeat in Depressive Pseudodementia. JAMA Neurol. 71, 775.

Bischof, J., Maeda, R.K., Hediger, M., Karch, F., Basler, K., 2007. An optimized transgenesis system for Drosophila using germ-line-specific C31 integrases. Proc. Natl. Acad. Sci. 104, 3312–3317.

Bjedov, I., Toivonen, J.M., Kerr, F., Slack, C., Jacobson, J., Foley, A., Partridge, L., 2010. Mechanisms of Life Span Extension by Rapamycin in the Fruit Fly Drosophila melanogaster. Cell Metab. 11, 35–46.

Blackstone, C., O'Kane, C.J., Reid, E., 2010. Hereditary spastic paraplegias: membrane traffic and the motor pathway. Nat. Rev. Neurosci. 12, 31–42.

Bocchetta, M., Cardoso, M.J., Cash, D.M., Ourselin, S., Warren, J.D., Rohrer, J.D., 2016. Patterns of regional cerebellar atrophy in genetic frontotemporal dementia. NeuroImage Clin. 11, 287–290.

Boeve, B.F., Boylan, K.B., Graff-Radford, N.R., Dejesus-Hernandez, M., Knopman, D.S., Pedraza, O., Vemuri, P., Jones, D., Lowe, V., Murray, M.E., Dickson, D.W., Josephs, K.A., Rush, B.K., Machulda, M.M., Fields, J.A., Ferman, T.J., Baker, M., Rutherford, N.J., Adamson, J., Wszolek, Z.K., Adeli, A., Savica, R., Boot, B., Kuntz, K.M., Gavrilova, R., Reeves, A., Whitwell, J., Kantarci, K., Jack, C.R., Parisi, J.E., Lucas, J.A., Petersen, R.C., Rademakers, R., 2012. Characterization of frontotemporal dementia and/or amyotrophic lateral sclerosis associated with the GGGGCC repeat expansion in C9ORF72. Brain 135, 765–783.

Boeynaems, S., Bogaert, E., Kovacs, D., Konijnenberg, A., Timmerman, E., Volkov, A., Guharoy, M., De Decker, M., Jaspers, T., Ryan, V.H., Janke, A.M., Baatsen, P., Vercruysse, T., Kolaitis, R.M., Daelemans, D., Taylor, J.P., Kedersha, N., Anderson, P., Impens, F., Sobott, F., Schymkowitz, J., Rousseau, F., Fawzi, N.L., Robberecht, W., Van Damme, P., Tompa, P., Van Den Bosch, L., 2017. Phase Separation of C9orf72 Dipeptide Repeats Perturbs Stress Granule Dynamics. Mol. Cell 65, 1044–1055.e5.

Boeynaems, S., Bogaert, E., Michiels, E., Gijselinck, I., Sieben, A., Jovičić, A., De Baets, G., Scheveneels, W., Steyaert, J., Cuijt, I., Verstrepen, K.J., Callaerts, P., Rousseau, F., Schymkowitz, J., Cruts, M., Van Broeckhoven, C., Van Damme, P., Gitler, A.D., Robberecht, W., Van Den Bosch, L., 2016a. Drosophila screen connects nuclear transport genes to DPR pathology in c9ALS/FTD. Sci. Rep. 6, 20877.

Boeynaems, S., Bogaert, E., Van Damme, P., Van Den Bosch, L., 2016b. Inside out: the role of nucleocytoplasmic transport in ALS and FTLD. Acta Neuropathol.

Bosco, D.A., Lemay, N., Ko, H.K., Zhou, H., Burke, C., Kwiatkowski, T.J., Sapp, P., Mckenna-Yasek, D., Brown, R.H., Hayward, L.J., 2010. Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules. Hum. Mol. Genet. 19, 4160–4175.

Boxer, A.L., Mackenzie, I.R., Boeve, B.F., Baker, M., Seeley, W.W., Crook, R., Feldman, H., Hsiung, G.-Y.R., Rutherford, N., Laluz, V., Whitwell, J., Foti, D., McDade, E., Molano, J., Karydas, A., Wojtas, A., Goldman, J., Mirsky, J., Sengdy, P., DeArmond, S., Miller, B.L., Rademakers, R., 2011. Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family. J. Neurol. Neurosurg. Psychiatry 82, 196–203.

Brangwynne, C.P., 2013. Phase transitions and size scaling of membrane-less organelles. J. Cell Biol. 203, 875–881.

Brangwynne, C.P., Tompa, P., Pappu, R. V., 2015. Polymer physics of intracellular phase transitions. Nat. Phys. 11, 899–904.

Branson, K., Robie, A.A., Bender, J., Perona, P., Dickinson, M.H., 2009. High-throughput ethomics in large groups of Drosophila. Nat. Methods 6, 451–457.

Brettschneider, J., Toledo, J.B., van Deerlin, V.M., Elman, L., McCluskey, L., Lee, V.M.Y., Trojanowski, J.Q., 2012. Microglial activation correlates with disease progression and upper motor neuron clinical symptoms in amyotrophic lateral sclerosis. PLoS One 7, e39216.

Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.P., Hudson, T., Sohn, R., Zemelman, B., Snell, R.G., Rundle, S.A., Crow, S., Davies, J., Shelbourne, P., Buxton, J., Jones, C., Juvonen, V., Johnson, K., Harper, P.S., Shaw, D.J., Housman, D.E., 1992. Molecular basis of myotonic dystrophy: Expansion of a trinucleotide

(CTG) repeat at the 3’ end of a transcript encoding a protein kinase family member. Cell 68, 799–808.

Bucci, C., Thomsen, P., Nicoziani, P., McCarthy, J., van Deurs, B., 2000. Rab7: a key to lysosome biogenesis. Mol. Biol. Cell 11, 467–80.

Buchan, J.R., Parker, R., 2009. Eukaryotic Stress Granules: The Ins and Outs of Translation. Mol. Cell.

Bunton-Stasyshyn, R.K.A., Saccon, R.A., Fratta, P., Fisher, E.M.C., 2015. SOD1 Function and Its Implications for Amyotrophic Lateral Sclerosis Pathology. Neurosci. 21, 519–529.

Burberry, A., Suzuki, N., Wang, J.-Y., Moccia, R., Mordes, D.A., Stewart, M.H., Suzuki-Uematsu, S., Ghosh, S., Singh, A., Merkle, F.T., Koszka, K., Li, Q.-Z., Zon, L., Rossi, D.J., Trowbridge, J.J., Notarangelo, L.D., Eggan, K., 2016. Loss-of-function mutations in the C9ORF72 mouse ortholog cause fatal autoimmune disease. Sci. Transl. Med. 8, 347ra93-347ra93.

Burguete, A.S., Almeida, S., Gao, F.B., Kalb, R., Akins, M.R., Bonini, N.M., 2015. GGGGCC microsatellite RNA is neuritically localized, induces branching defects, and perturbs transport granule function. Elife 4, e08881.

Calderwood, S.K., Murshid, A., Prince, T., 2009. The shock of aging: Molecular chaperones and the heat shock response in longevity and aging - A mini-review. Gerontology.

Callister, J.B., Ryan, S., Sim, J., Rollinson, S., Pickering-Brown, S.M., 2016. Modelling C9orf72 dipeptide repeat proteins of a physiologically relevant size. Hum. Mol. Genet. 25, 5069–5082.

Calvo, A., Moglia, C., Canosa, A., Cistaro, A., Valentini, C., Carrara, G., Soldano, E., Ilardi, A., Bersano, E., Bertuzzo, D., Brunetti, M., Ossola, I., Restagno, G., Chio, A., 2012. Amyotrophic lateral sclerosis/frontotemporal dementia with predominant manifestations of obsessive-compulsive disorder associated to GGGGCC expansion of the c9orf72 gene. J. Neurol.

Carazo-Salas, R.E., Guarguaglini, G., Gruss, O.J., Segref, A., Karsenti, E., Mattaj, I.W., 1999. Generation of GTP-bound Ran by RCC1 is required for chromatin-induced mitotic spindle formation. Nature 400, 178–81.

Cautain, B., Hill, R., De Pedro, N., Link, W., 2015. Components and regulation of nuclear transport processes. FEBS J.

Celona, B., Von Dollen, J., Vatsavayai, S.C., Kashima, R., Johnson, J.R., Tang, A.A., Hata, A., Miller, B.L., Huang, E.J., Krogan, N.J., Seeley, W.W., Black, B.L., 2017. Suppression of c9orf72 RNA repeat-induced neurotoxicity by the ALS-associated RNA-binding protein Zfp106. Elife 6.

Chang, W.L., Tarn, W.Y., 2009. A role for transportin in deposition of TTP to cytoplasmic RNA granules and mRNA decay. Nucleic Acids Res. 37, 6600–6612.

Chang, Y.J., Jeng, U.S., Chiang, Y.L., Hwang, I.S., Chen, Y.R., 2016. The glycine-alanine dipeptide repeat from C9orf72 hexanucleotide expansions forms toxic amyloids possessing cell-to-cell transmission properties. J. Biol. Chem. 291, 4903–4911.

Chen, Y., Lin, Z., Chen, X., Cao, B., Wei, Q., Ou, R., Zhao, B., Song, W., Wu, Y., Shang, H.F., 2015. Large C9orf72 repeat expansions are seen in Chinese patients with sporadic amyotrophic lateral sclerosis. Neurobiol. Aging 38, 217.e15-217.e22.

Cherfils, J., Zeghouf, M., 2013. Regulation of Small GTPases by GEFs, GAPs, and GDIs. Physiol. Rev. 93, 269–309.

Chew, J., Gendron, T.F., Prudencio, M., Sasaguri, H., Zhang, Y.-J., Castanedes-Casey, M., Lee, C.W., Jansen-West, K., Kurti, A., Murray, M.E., Bieniek, K.F., Bauer, P.O., Whitelaw, E.C., Rousseau, L., Stankowski, J.N., Stetler, C., Daughrity, L.M., Perkerson, E.A., Desaro, P., Johnston, A., Overstreet, K., Edbauer, D., Rademakers, R., Boylan, K.B., Dickson, D.W., Fryer, J.D., Petrucelli, L., 2015. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. Science (80-.). 348, 1151–1154.

Chiang, H.-H., Forsell, C., Lindström, A.-K., Lilius, L., Thonberg, H., Nennesmo, I., Graff, C., 2017. No common founder for C9orf72 expansion mutation in Sweden. J. Hum. Genet. 62, 321–324.

Chiò, A., Calvo, A., Moglia, C., Mazzini, L., Mora, G., 2011. Phenotypic heterogeneity of amyotrophic lateral sclerosis: a population based study. J. Neurol. Neurosurg. Psychiatry 82, 740–746.

Chiò, A., Borghero, G., Restagno, G., Mora, G., Drepper, C., Traynor, B.J., Sendtner, M., Brunetti, M., Ossola, I., Calvo, A., Pugliatti, M., Sotgiu, M.A., Murru, M.R., Marrosu, M.G., Marrosu, F., Marinou, K., Mandrioli, J., Sola, P., Caponnetto, C., Mancardi, G., Mandich, P., La Bella, V., Spataro, R., Conte, A., Monsurrò, M.R., Tedeschi, G., Pisano, F., Bartolomei, I., Salvi, F., Lauria Pinter, G., Simone, I., Logroscino, G., Gambardella, A., Quattrone, A., Lunetta, C., Volanti, P., Zollino, M., Penco, S., Battistini, S., Renton, A.E., Majounie, E., Abramzon, Y., Conforti, F.L., Giannini, F., Corbo, M., Sabatelli, M., Moglia, C., Cammarosano, S., Fuda, G., Canosa, A., Gallo, S., Papetti, L., Luigetti, M., Lattante, S., Marangi, G., Colletti, T., Ricci, C., Origone, P., Floris, G., Cannas, A., Piras, V., Parish, L.D., Solinas, G., Ulgheri, L., Ticca, A., Izzo, F., Laiola, A., Trojsi, F., 2012a. Clinical characteristics of patients with familial amyotrophic lateral sclerosis carrying the pathogenic GGGGCC hexanucleotide repeat expansion of C9ORF72. Brain 135, 784–793.

Chiò, A., Calvo, A., Mazzini, L., Cantello, R., Mora, G., Moglia, C., Corrado, L., D'Alfonso, S., Majounie, E., Renton, A., Pisano, F., Ossola, I., Brunetti, M., Traynor, B.J., Restagno, G., 2012b. Extensive genetics of ALS : A population-based study in Italy. Neurology 79, 1983–1989.

Chiò, A., Restagno, G., Brunetti, M., Ossola, I., Calvo, A., Mora, G., Sabatelli, M., Monsurrò, M.R., Battistini, S., Mandrioli, J., Salvi, F., Spataro, R., Schymick, J., Traynor, B.J., La Bella, V., 2009. Two Italian kindreds with familial amyotrophic lateral sclerosis due to FUS mutation. Neurobiol. Aging 30, 1272–1275.

Cirulli, E.T., Lasseigne, B.N., Petrovski, S., Sapp, P.C., Dion, P.A., Leblond, C.S., Couthouis, J., Lu, Y.-F., Wang, Q., Krueger, B.J., Ren, Z., Keebler, J., Han, Y., Levy, S.E., Boone, B.E., Wimbish, J.R., Waite, L.L., Jones, A.L., Carulli, J.P., Day-Williams, A.G., Staropoli, J.F., Xin, W.W., Chesi, A., Raphael, A.R., McKenna-Yasek, D., Cady, J., Vianney de Jong, J.M.B., Kenna, K.P., Smith, B.N., Topp, S., Miller, J., Gkazi, A., Al-Chalabi, A., van den Berg, L.H., Veldink, J., Silani, V., Ticozzi, N., Shaw, C.E., Baloh, R.H., Appel, S., Simpson, E., Lagier-Tourenne, C., Pulst, S.M., Gibson, S., Trojanowski, J.Q., Elman, L., McCluskey, L., Grossman, M., Shneider, N.A., Chung, W.K., Ravits, J.M., Glass, J.D., Sims, K.B., Van Deerlin, V.M., Maniatis, T., Hayes, S.D., Ordureau, A., Swarup, S., Landers, J., Baas, F., Allen, A.S., Bedlack, R.S., Harper, J.W., Gitler, A.D., Rouleau, G.A., Brown, R., Harms, M.B., Cooper, G.M., Harris, T., Myers, R.M., Goldstein, D.B., 2015. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. Science (80-.). 347, 1436–1441.

Ciura, S., Lattante, S., Le Ber, I., Latouche, M., Tostivint, H., Brice, A., Kabashi, E., 2013. Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. Ann. Neurol. 74, 180–187.

Cleary, E.M., Pal, S., Azam, T., Moore, D.J., Swingler, R., Gorrie, G., Stephenson, L., Colville, S., Chandran, S., Porteous, M., Warner, J.P., 2016. Improved PCR based methods for detecting C9orf72 hexanucleotide repeat expansions. Mol. Cell. Probes 30, 218–224.

Cleary, J.D., Ranum, L.P., 2017. New developments in RAN translation: insights from multiple diseases. Curr. Opin. Genet. Dev.

Colley, N.J., 2012. Retinal degeneration in the fly. In: Advances in Experimental Medicine and Biology. NIH Public Access, pp. 407–414.

Colombrita, C., Zennaro, E., Fallini, C., Weber, M., Sommacal, A., Buratti, E., Silani, V., Ratti, A., 2009. TDP-43 is recruited to stress granules in conditions of oxidative insult. J. Neurochem. 111, 1051–1061.

Conicella, A.E., Zerze, G.H., Mittal, J., Fawzi, N.L., 2016. ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. Structure 24, 1537–1549.

Conlon, E.G., Lu, L., Sharma, A., Yamazaki, T., Tang, T., Shneider, N.A., Manley, J.L., 2016. The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. Elife 5.

Cooper-Knock, J., Hewitt, C., Highley, J.R., Brockington, A., Milano, A., Man, S., Martindale, J., Hartley, J., Walsh, T., Gelsthorpe, C., Baxter, L., Forster, G., Fox, M., Bury, J., Mok, K., McDermott, C.J., Traynor, B.J., Kirby, J., Wharton, S.B., Ince, P.G., Hardy, J., Shaw, P.J., 2012. Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. Brain 135, 751–764.

Cooper-Knock, J., Higginbottom, A., Stopford, M.J., Highley, J.R., Ince, P.G., Wharton, S.B., Pickering-Brown, S., Kirby, J., Hautbergue, G.M., Shaw, P.J., 2015. Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy. Acta Neuropathol. 130, 63–75.

Cooper-Knock, J., Walsh, M.J., Higginbottom, A., Highley, J.R., Dickman, M.J., Edbauer, D., Ince, P.G., Wharton, S.B., Wilson, S.A., Kirby, J., Hautbergue, G.M., Shaw, P.J., 2014. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. Brain 137, 2040–2051.

Corcia, P., Vourc'h, P., Guennoc, A.-M., Del Mar Amador, M., Blasco, H., Andres, C., Couratier, P., Gordon, P.H., Meininger, V., 2015. Pure cerebellar ataxia linked to large C9orf72 repeat expansion. Amyotroph. Lateral Scler. Front. Degener. 8421, 1–3.

Coulom, H., 2004. Chronic Exposure to Rotenone Models Sporadic Parkinson's Disease in Drosophila melanogaster. J. Neurosci. 24, 10993–10998.

Coyne, A.N., Siddegowda, B.B., Estes, P.S., Johannesmeyer, J., Kovalik, T., Daniel, S.G., Pearson, A., Bowser, R., Daniela, X., Zarnescu, C., 2014. Neurobiology of Disease Futsch/MAP1B mRNA Is a Translational Target of TDP-43 and Is Neuroprotective in a Drosophila Model of Amyotrophic Lateral Sclerosis. J. Neurosci. 34, 15962–74.

Crenshaw, E., Leung, B.P., Kwok, C.K., Sharoni, M., Olson, K., Sebastian, N.P., Ansaloni, S., Schweitzer-Stenner, R., Akins, M.R., Bevilacqua, P.C., Saunders, A.J., 2015. Amyloid precursor protein translation is regulated by a 3'UTR guanine quadruplex. PLoS One 10, e0143160.

Cristofani, R., Crippa, V., Vezzoli, G., Rusmini, P., Galbiati, M., Cicardi, M.E., Meroni, M., Ferrari, V., Tedesco, B., Piccolella, M., Messi, E., Carra, S., Poletti, A., 2017. The small heat shock protein B8 (HSPB8) efficiently removes aggregating species of dipeptides produced in C9ORF72-related neurodegenerative diseases. Cell Stress Chaperones 1–12.

Cruts, M., Engelborghs, S., van der Zee, J., Van Broeckhoven, C., 2015. C9orf72-Related Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. In: GeneReviews(R). University of Washington, Seattle.

Cruts, M., Gijselinck, I., Van Langenhove, T., van der Zee, J., Van Broeckhoven, C., 2013. Current insights into the C9orf72 repeat expansion diseases of the FTLD/ALS spectrum. Trends Neurosci.

Dafinca, R., Scaber, J., Ababneh, N., Lalic, T., Weir, G., Christian, H., Vowles, J., Douglas, A.G.L., Fletcher-Jones, A., Browne, C., Nakanishi, M., Turner, M.R., Wade-Martins, R., Cowley, S.A., Talbot, K., 2016. C9orf72 Hexanucleotide Expansions Are Associated with Altered Endoplasmic Reticulum Calcium Homeostasis and Stress Granule Formation in Induced Pluripotent Stem Cell-Derived Neurons from Patients with Amyotrophic Lateral Sclerosis and Frontotemporal Demen. Stem Cells 34, 2063–2078.

D'amico, E., Pasmantier, M., Lee, Y.W., Weimer, L., Mitsumoto, H., 2013. Clinical evolution of pure upper motor neuron disease/dysfunction (PUMMD). Muscle and Nerve 47, 28–32.

D'Angelo, M.A., Raices, M., Panowski, S.H., Hetzer, M.W., 2009. Age-Dependent Deterioration of Nuclear Pore Complexes Causes a Loss of Nuclear Integrity in Postmitotic Cells. Cell 136, 284–295.

David, D.C., Ollikainen, N., Trinidad, J.C., Cary, M.P., Burlingame, A.L., Kenyon, C., 2010. Widespread protein aggregation as an inherent part of aging in C. elegans. PLoS Biol. 8, 47–48.

Davidson, Y., Robinson, A.C., Liu, X., Wu, D., Troakes, C., Rollinson, S., Masuda-Suzukake, M., Suzuki, G., Nonaka, T., Shi, J., Tian, J., Hamdalla, H., Ealing, J., Richardson, A., Jones, M., Pickering-Brown, S., Snowden, J.S., Hasegawa, M., Mann, D.M.A., 2016. Neurodegeneration in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9orf72 is linked to TDP-43 pathology and not associated with aggregated forms of dipeptide repeat proteins. Neuropathol. Appl. Neurobiol. 42, 242–254.

Davidson, Y.S., Barker, H., Robinson, A.C., Thompson, J.C., Harris, J., Troakes, C., Smith, B., Al-Saraj, S., Shaw, C., Rollinson, S., Masuda-Suzukake, M., Hasegawa, M., Pickering-Brown, S., Snowden, J.S., Mann, D.M., 2014. Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. Acta Neuropathol. Commun. 2, 70.

de Haro, M., Al-Ramahi, I., De Gouyon, B., Ukani, L., Rosa, A., Faustino, N.A., Ashizawa, T., Cooper, T.A., Botas, J., 2006. MBNL1 and CUGBP1 modify expanded CUG-induced toxicity in a Drosophila model of myotonic dystrophy type 1. Hum. Mol. Genet. 15, 2138–2145.

DeJesus-Hernandez, M., Finch, N.C.A., Wang, X., Gendron, T.F., Bieniek, K.F., Heckman, M.G., Vasilevich, A., Murray, M.E., Rousseau, L., Weesner, R., Lucido, A., Parsons, M., Chew, J., Josephs, K.A., Parisi, J.E., Knopman, D.S., Petersen, R.C., Boeve, B.F., Graff-Radford, N.R., de Boer, J., Asmann, Y.W., Petrucelli, L., Boylan, K.B., Dickson, D.W., van Blitterswijk, M., Rademakers, R.,

2017. In-depth clinico-pathological examination of RNA foci in a large cohort of C9ORF72 expansion carriers. Acta Neuropathol. 134, 255–269.

DeJesus-Hernandez, M., Kocerha, J., Finch, N., Crook, R., Baker, M., Desaro, P., Johnston, A., Rutherford, N., Wojtas, A., Kennelly, K., Wszolek, Z.K., Graff-Radford, N., Boylan, K., Rademakers, R., 2010. De novo truncating FUS gene mutation as a cause of sporadic amyotrophic lateral sclerosis. Hum. Mutat. 31, E1377–E1389.

DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.C.A., Flynn, H., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsiung, G.Y.R., Karydas, A., Seeley, W.W., Josephs, K.A., Coppola, G., Geschwind, D.H., Wszolek, Z.K., Feldman, H., Knopman, D.S., Petersen, R.C., Miller, B.L., Dickson, D.W., Boylan, K.B., Graff-Radford, N.R., Rademakers, R., 2011. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. Neuron 72, 245–256.

Dewey, C.M., Cenik, B., Sephton, C.F., Dries, D.R., Mayer, P., Good, S.K., Johnson, B.A., Herz, J., Yu, G., 2011. TDP-43 Is Directed to Stress Granules by Sorbitol, a Novel Physiological Osmotic and Oxidative Stressor. Mol. Cell. Biol. 31, 1098–1108.

Diaper, D.C., Adachi, Y., Sutcliffe, B., Humphrey, D.M., Elliott, C.J.H., Stepto, A., Ludlow, Z.N., Broeck, L. Vanden, Callaerts, P., Dermaut, B., Al-Chalabi, A., Shaw, C.E., Robinson, I.M., Hirth, F., 2013a. Loss and gain of Drosophila TDP-43 impair synaptic efficacy and motor control leading to age-related neurodegeneration by loss-of-function phenotypes. Hum. Mol. Genet. 22, 1539–1557.

Diaper, D.C., Adachi, Y., Lazarou, L., Greenstein, M., Simoes, F.A., Di Domenico, A., Solomon, D.A., Lowe, S., Alsubaie, R., Cheng, D., Buckley, S., Humphrey, D.M., Shaw, C.E., Hirth, F., 2013b. Drosophila TDP-43 dysfunction in glia and muscle cells cause cytological and behavioural phenotypes that characterize ALS and FTLD. Hum. Mol. Genet. 22, 3883–3893.

Dols-Icardo, O., García-Redondo, A., Rojas-García, R., Sánchez-Valle, R., Noguera, A., Gó mez-Tortosa, E., Pastor, P., Hernández, I., Esteban-Pérez, J., Suárez-Calvet, M., Antó n-Aguirre, S., Amer, G., Ortega-Cubero, S., Blesa, R., Fortea, J., Alcolea, D., Capdevila, A., Antonell, A., Lladó, A., Munoz-Blanco, J.L., Mora, J.S., LucíaGalán-Dá vila, De Rivera, F.J.R., Lleó, A., Clarimón, J., 2014. Characterization of the repeat expansion size in C9orf72 in amyotrophic lateral sclerosis and frontotemporal dementia. Hum. Mol. Genet. 23, 749–754.

Donnelly, C.J., Zhang, P.W., Pham, J.T., Heusler, A.R., Mistry, N.A., Vidensky, S., Daley, E.L., Poth, E.M., Hoover, B., Fines, D.M., Maragakis, N., Tienari, P.J., Petrucelli, L., Traynor, B.J., Wang, J., Rigo, F., Bennett, C.F., Blackshaw, S., Sattler, R., Rothstein, J.D., 2013. RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. Neuron 80, 415–428.

Dormann, D., Haass, C., 2011. TDP-43 and FUS: A nuclear affair. Trends Neurosci.

Dormann, D., Madl, T., Valori, C.F., Bentmann, E., Tahirovic, S., Abou-Ajram, C., Kremmer, E., Ansorge, O., Mackenzie, I.R.A., Neumann, M., Haass, C., 2012. Arginine methylation next to the PY-NLS modulates Transportin binding and nuclear import of FUS. EMBO J. 31, 4258–4275.

Dormann, D., Rodde, R., Edbauer, D., Bentmann, E., Fischer, I., Hruscha, A., Than, M.E., Mackenzie, I.R.A., Capell, A., Schmid, B., Neumann, M., Haass, C., 2010. ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. EMBO J. 29, 2841–2857.

Düchler, M., 2012. G-quadruplexes: targets and tools in anticancer drug design. J. Drug Target. 20, 389–400.

Edbauer, D., Haass, C., 2016. An amyloid-like cascade hypothesis for C9orf72 ALS/FTD. Curr. Opin. Neurobiol.

Ederle, H., Dormann, D., 2017. TDP-43 and FUS en route from the nucleus to the cytoplasm. FEBS Lett.

Elden, A.C., Kim, H.-J., Hart, M.P., Chen-Plotkin, A.S., Johnson, B.S., Fang, X., Armakola, M., Geser, F., Greene, R., Lu, M.M., Padmanabhan, A., Clay-Falcone, D., McCluskey, L., Elman, L., Juhr, D., Gruber, P.J., Rüb, U., Auburger, G., Trojanowski, J.Q., Lee, V.M.-Y., Van Deerlin, V.M., Bonini, N.M., Gitler, A.D., 2010. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. Nature 466, 1069–1075.

Es, M.A. Van, Veldink, J.H., Saris, C.G.J., Blauw, H.M., Vught, P.W.J. Van, Birve, A., Lemmens, R., Schelhaas, H.J., Groen, E.J.N., Huisman, M.H.B., Kooi, A.J. Van Der, Visser, D., Dahlberg, C., Estrada, K., Rivadeneira, F., Hofman, A., Zwarts, M.J., Doormaal, P.T.C. Van, Rujescu, D., Strengman, E., Giegling, I., Muglia, P., Tomik, B., Slowik, A., Uitterlinden, A.G., Hendrich, C., Waibel, S., Meyer, T., Ludolph, A.C., Glass, J.D., Purcell, S., Cichon, S., Nöthen, M.M., Wokke, J.H.J., Cronin, S., Mclaughlin, R.L., Hardiman, O., Fumoto, K., van Es, M.A., Veldink, J.H., Saris, C.G.J., Blauw, H.M., van Vught, P.W.J., Birve, A., Lemmens, R., Schelhaas, H.J., Groen, E.J.N., Huisman, M.H.B., van der Kooi, A.J., de Visser, M., Dahlberg, C., Estrada, K., Rivadeneira, F., Hofman, A., Zwarts, M.J., van Doormaal, P.T.C., Rujescu, D., Strengman, E., Giegling, I., Muglia, P., Tomik, B., Slowik, A., Uitterlinden, A.G., Hendrich, C., Waibel, S., Meyer, T., Ludolph, A.C., Glass, J.D., Purcell, S., Cichon, S., Nöthen, M.M., Wichmann, H.-E., Schreiber, S., Vermeulen, S.H.H.M., Kiemeney, L.A., Wokke, J.H.J., Cronin, S., Mclaughlin, R.L., Hardiman, O., Fumoto, K., Pasterkamp, R.J., Meininger, V., Melki, J., Leigh, P.N., Shaw, C.E., Landers, J.E., Al-Chalabi, A., Brown, R.H., Robberecht, W., Andersen, P.M., Ophoff, R.A., van den Berg, L.H., 2009. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. Nat. Genet. 41, 1083–1087.

Eskelinen, E.-L., Saftig, P., 2009. Autophagy: A lysosomal degradation pathway with a central role in health and disease. Biochim. Biophys. Acta - Mol. Cell Res. 1793, 664–673.

Evers, M.M., Toonen, L.J.A., van Roon-Mom, W.M.C., 2015. Antisense oligonucleotides in therapy for neurodegenerative disorders. Adv. Drug Deliv. Rev.

Fagerlund, R., Kinnunen, L., Köhler, M., Julkunen, I., Melén, K., 2005. NF-kB is transported into the nucleus by importin α3 and importin α4. J. Biol. Chem. 280, 15942–15951.

Faravelli, I., Nizzardo, M., Comi, G.P., Corti, S., 2015. Spinal muscular atrophy—recent therapeutic advances for an old challenge. Nat. Rev. Neurol. 11, 351–359.

Fardaei, M., Larkin, K., Brook, J.D., Hamshere, M.G., 2001. In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. Nucleic Acids Res. 29, 2766–71.

Farg, M.A., Konopka, A., Soo, K.Y., Ito, D., Atkin, J.D., 2017. The DNA damage response (DDR) is induced by the C9orf72 repeat expansion in amyotrophic lateral sclerosis. Hum. Mol. Genet. 26, 2882–2896.

Farg, M.A., Sundaramoorthy, V., Sultana, J.M., Yang, S., Atkinson, R.A.K., Levina, V., Halloran, M.A., Gleeson, P.A., Blair, I.P., Soo, K.Y., King, A.E., Atkin, J.D., 2014. C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. Hum. Mol. Genet. 23, 3579–3595.

Flores, B.N., Dulchavsky, M.E., Krans, A., Sawaya, M.R., Paulson, H.L., Todd, P.K., Barmada, S.J., Ivanova, M.I., 2016. Distinct c9orf72-associated dipeptide repeat structures correlate with neuronal toxicity. PLoS One 11, e0165084.

Floyd, A.G., Yu, Q.P., Piboolnurak, P., Tang, M.X., Fang, Y., Smith, W.A., Yim, J., Rowland, L.P., Mitsumoto, H., Pullman, S.L., 2009. Transcranial magnetic stimulation in ALS: Utility of central motor conduction tests. Neurology 72, 498–504.

Flynn, K.C., Hellal, F., Neukirchen, D., Jacob, S., Tahirovic, S., Dupraz, S., Stern, S., Garvalov, B.K., Gurniak, C., Shaw, A.E., Meyn, L., Wedlich-Söldner, R., Bamburg, J.R., Small, J.V., Witke, W., Bradke, F., 2012. ADF/Cofilin-Mediated Actin Retrograde Flow Directs Neurite Formation in the Developing Brain. Neuron 76, 1091–1107.

Frakes, A.E., Ferraiuolo, L., Haidet-Phillips, A.M., Schmelzer, L., Braun, L., Miranda, C.J., Ladner, K.J., Bevan, A.K., Foust, K.D., Godbout, J.P., Popovich, P.G., Guttridge, D.C., Kaspar, B.K., 2014. Microglia induce motor neuron death via the classical NF-κB pathway in amyotrophic lateral sclerosis. Neuron 81, 1009–1023.

Frasch, M., 1991. The maternally expressed Drosophila gene encoding the chromatin-binding protein BJ1 is a homolog of the vertebrate gene Regulator of Chromatin Condensation, RCC1. EMBO J. 10, 1225–36.

Fratta, P., Mizielinska, S., Nicoll, A.J., Zloh, M., Fisher, E.M.C., Parkinson, G., Isaacs, A.M., 2012. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. Sci. Rep. 2, 1016.

Fratta, P., Polke, J.M., Newcombe, J., Mizielinska, S., Lashley, T., Poulter, M., Beck, J., Preza, E., Devoy, A., Sidle, K., Howard, R., Malaspina, A., Orrell, R.W., Clarke, J., Lu, C.H., Mok, K., Collins, T., Shoaii, M., Nanji, T., Wray, S., Adamson, G., Pittman, A., Renton, A.E., Traynor, B.J., Sweeney, M.G., Revesz, T., Houlden, H., Mead, S., Isaacs, A.M., Fisher, E.M.C., 2015. Screening a UK amyotrophic lateral sclerosis cohort provides evidence of multiple origins of the C9orf72 expansion. Neurobiol. Aging 36, 546.e1-546.e7.

Fratta, P., Poulter, M., Lashley, T., Rohrer, J.D., Polke, J.M., Beck, J., Ryan, N., Hensman, D., Mizielinska, S., Waite, A.J., Lai, M.C., Gendron, T.F., Petrucelli, L., Fisher, E.M.C., Revesz, T., Warren, J.D., Collinge, J., Isaacs, A.M., Mead, S., 2013. Homozygosity for the C9orf72 GGGGCC repeat expansion in frontotemporal dementia. Acta Neuropathol. 126, 401–409.

Freibaum, B.D., Chitta, R.K., High, A.A., Taylor, J.P., 2010. Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. J. Proteome Res. 9, 1104–1120.

Freibaum, B.D., Lu, Y., Lopez-Gonzalez, R., Kim, N.C., Almeida, S., Lee, K.-H., Badders, N., Valentine, M., Miller, B.L., Wong, P.C., Petrucelli, L., Kim, H.J., Gao, F.-B., Taylor, J.P., 2015. GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. Nature 525, 129–133.

Freibaum, B.D., Taylor, J.P., 2017. The Role of Dipeptide Repeats in C9ORF72-Related ALS-FTD. Front. Mol. Neurosci. 10, 35.

Freischmidt, A., Wieland, T., Richter, B., Ruf, W., Schaeffer, V., Müller, K., Marroquin, N., Nordin, F., Hübers, A., Weydt, P., Pinto, S., Press, R., Millecamps, S., Molko, N., Bernard, E., Desnuelle, C., Soriani, M.-H., Dorst, J., Graf, E., Nordström, U., Feiler, M.S., Putz, S., Boeckers, T.M., Meyer, T., Winkler, A.S., Winkelman, J., de Carvalho, M., Thal, D.R., Otto, M., Brännström, T., Volk, A.E., Kursula, P., Danzer, K.M., Lichtner, P., Dikic, I., Meitinger, T., Ludolph, A.C., Strom, T.M., Andersen, P.M., Weishaupt, J.H., 2015. Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. Nat. Neurosci. 18, 631–636.

Frey, S., Görlich, D., 2007. A Saturated FG-Repeat Hydrogel Can Reproduce the Permeability Properties of Nuclear Pore Complexes. Cell 130, 512–523.

Fu, X.-D., Ares, M., 2014. Context-dependent control of alternative splicing by RNA-binding proteins. Nat. Rev. Genet. 15, 689–701.

Fujimura, K., Suzuki, T., Yasuda, Y., Murata, M., Katahira, J., Yoneda, Y., 2010. Identification of importin ??1 as a novel constituent of RNA stress granules. Biochim. Biophys. Acta - Mol. Cell Res. 1803, 865–871.

Galimberti, D., Reif, A., Dell'Osso, B., Kittel-Schneider, S., Leonhard, C., Herr, A., Palazzo, C., Villa, C., Fenoglio, C., Serpente, M., Cioffi, S.M.G., Prunas, C., Paoli, R.A., Altamura, A.C., Scarpini, E., 2016. The C9ORF72 hexanucleotide repeat expansion is a rare cause of schizophrenia. Neurobiol. Aging 35, 1214.e7-1214.e10.

Gami, P., Murray, C., Schottlaender, L., Bettencourt, C., De Pablo Fernandez, E., Mudanohwo, E., Mizielinska, S., Polke, J.M., Holton, J.L., Isaacs, A.M., Houlden, H., Revesz, T., Lashley, T., 2015. A 30-unit hexanucleotide repeat expansion in C9orf72 induces pathological lesions with dipeptide-repeat proteins and RNA foci, but not TDP-43 inclusions and clinical disease. Acta Neuropathol.

Gao, F., Almeida, S., Lopez-Gonzalez, R., 2017. Dysregulated molecular pathways in amyotrophic lateral sclerosis–frontotemporal dementia spectrum disorder. EMBO J. 36, e201797568.

García-Redondo, A., Dols-Icardo, O., Rojas-García, R., Esteban-Pérez, J., Cordero-Vázquez, P., Muñoz-Blanco, J.L., Catalina, I., González-Muñoz, M., Varona, L., Sarasola, E., Povedano, M., Sevilla, T., Guerrero, A., Pardo, J., de Munain, A.L., Márquez-Infante, C., de Rivera, F.J.R., Pastor, P., Jericó, I., de Arcaya, A.Á., Mora, J.S., Clarimón, J., 2013. Analysis of the C9orf72 Gene in Patients with Amyotrophic Lateral Sclerosis in Spain and Different Populations Worldwide. Hum. Mutat. 34, 79–82.

Gautier, G., Verschueren, A., Monnier, A., Attarian, S., Salort-Campana, E., Pouget, J., 2010. ALS with respiratory onset: Clinical features and effects of non-invasive ventilation on the prognosis. Amyotroph. Lateral Scler. 11, 379–382.

Gendron, T.F., Bieniek, K.F., Zhang, Y.J., Jansen-West, K., Ash, P.E.A., Caulfield, T., Daughrity, L., Dunmore, J.H., Castanedes-Casey, M., Chew, J., Cosio, D.M., Van Blitterswijk, M., Lee, W.C., Rademakers, R., Boylan, K.B., Dickson, D.W., Petrucelli, L., 2013. Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. Acta Neuropathol. 126, 829–844.

Gendron, T.F., van Blitterswijk, M., Bieniek, K.F., Daughrity, L.M., Jiang, J., Rush, B.K., Pedraza, O., Lucas, J.A., Murray, M.E., Desaro, P., Robertson, A., Overstreet, K., Thomas, C.S., Crook, J.E., Castanedes-Casey, M., Rousseau, L., Josephs, K.A., Parisi, J.E., Knopman, D.S., Petersen, R.C., Boeve, B.F., Graff-Radford, N.R., Rademakers, R., Lagier-Tourenne, C., Edbauer, D., Cleveland, D.W., Dickson, D.W., Petrucelli, L., Boylan, K.B., 2015. Cerebellar c9RAN proteins associate with clinical and neuropathological characteristics of C9ORF72 repeat expansion carriers. Acta Neuropathol. 130, 559–573.

Gendron, T.F., Chew, J., Stankowski, J.N., Hayes, L.R., Zhang, Y.-J., Prudencio, M., Carlomagno, Y., Daughrity, L.M., Jansen-West, K., Perkerson, E.A., O'Raw, A.,

Cook, C., Pregent, L., Belzil, V., van Blitterswijk, M., Tabassian, L.J., Lee, C.W., Yue, M., Tong, J., Song, Y., Castanedes-Casey, M., Rousseau, L., Phillips, V., Dickson, D.W., Rademakers, R., Fryer, J.D., Rush, B.K., Pedraza, O., Caputo, A.M., Desaro, P., Palmucci, C., Robertson, A., Heckman, M.G., Diehl, N.N., Wiggs, E., Tierney, M., Braun, L., Farren, J., Lacomis, D., Ladha, S., Fournier, C.N., McCluskey, L.F., Elman, L.B., Toledo, J.B., McBride, J.D., Tiloca, C., Morelli, C., Poletti, B., Solca, F., Prelle, A., Wuu, J., Jockel-Balsarotti, J., Rigo, F., Ambrose, C., Datta, A., Yang, W., Raitcheva, D., Antognetti, G., McCampbell, A., Van Swieten, J.C., Miller, B.L., Boxer, A.L., Brown, R.H., Bowser, R., Miller, T.M., Trojanowski, J.Q., Grossman, M., Berry, J.D., Hu, W.T., Ratti, A., Traynor, B.J., Disney, M.D., Benatar, M., Silani, V., Glass, J.D., Floeter, M.K., Rothstein, J.D., Boylan, K.B., Petrucelli, L., 2017. Poly(GP) proteins are a useful pharmacodynamic marker for C9ORF72 -associated amyotrophic lateral sclerosis. Sci. Transl. Med. 9, eaai7866.

Ghosh, S., Lippa, C.F., 2015. Clinical Subtypes of Frontotemporal Dementia. Am. J. Alzheimer's Dis. Other Dementiasr 30, 653–661.

Gijselinck, I., Van Mossevelde, S., van der Zee, J., Sieben, A., Engelborghs, S., De Bleecker, J., Ivanoiu, A., Deryck, O., Edbauer, D., Zhang, M., Heeman, B., Bäumer, V., Van den Broeck, M., Mattheijssens, M., Peeters, K., Rogaeva, E., De Jonghe, P., Cras, P., Martin, J.-J., de Deyn, P.P., Cruts, M., Van Broeckhoven, C., 2016. The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. Mol. Psychiatry 21, 1112–1124.

Gijselinck, I., Van Langenhove, T., van der Zee, J., Sleegers, K., Philtjens, S., Kleinberger, G., Janssens, J., Bettens, K., Van Cauwenberghe, C., Pereson, S., Engelborghs, S., Sieben, A., De Jonghe, P., Vandenberghe, R., Santens, P., De Bleecker, J., Maes, G., Bäumer, V., Dillen, L., Joris, G., Cuijt, I., Corsmit, E., Elinck, E., Van Dongen, J., Vermeulen, S., Van den Broeck, M., Vaerenberg, C., Mattheijssens, M., Peeters, K., Robberecht, W., Cras, P., Martin, J.J., De Deyn, P.P., Cruts, M., Van Broeckhoven, C., 2012. A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: A gene identification study. Lancet Neurol. 11, 54–65.

Gitler, A.D., Tsuiji, H., 2016. There has been an awakening: Emerging mechanisms of C9orf72 mutations in FTD/ALS. Brain Res.

Goldman, J.S., Quinzii, C., Dunning-Broadbent, J., Waters, C., Mitsumoto, H., Brannagan, T.H., Cosentino, S., Huey, E.D., Nagy, P., Kuo, S.-H., 2014. Multiple System Atrophy and Amyotrophic Lateral Sclerosis in a Family With Hexanucleotide Repeat Expansions in C9orf72. JAMA Neurol. 71, 771.

Gomez-Deza, J., Lee, Y., Troakes, C., Nolan, M., Al-Sarraj, S., Gallo, J.-M., Shaw, C.E., 2015. Dipeptide repeat protein inclusions are rare in the spinal cord and almost absent from motor neurons in C9ORF72 mutant amyotrophic lateral sclerosis and are unlikely to cause their degeneration. Acta Neuropathol. Commun. 3, 38.

Gómez-Tortosa, E., Gallego, J., Guerrero-Lopez, R., Marcos, A., Gil-Neciga, E., Jose Sainz, M., D??az, A., Franco-Macias, E., Trujillo-Tiebas, M.J., Ayuso, C., Perez-Perez, J., 2013. C9ORF72 hexanucleotide expansions of 20-22 repeats are associated with frontotemporal deterioration. Neurology 80, 366–370.

Gómez-Tortosa, E., Serrano, S., De Toledo, M., Perez-Perez, J., Sainz, M.J., 2014. Familial benign frontotemporal deterioration with C9ORF72 hexanucleotide expansion. Alzheimer's Dement. 10, S284–S289.

Gopal, P.P., Nirschl, J.J., Klinman, E., Holzbaur, E.L.F., 2017. Amyotrophic lateral sclerosis-linked mutations increase the viscosity of liquid-like TDP-43 RNP granules in neurons. Proc. Natl. Acad. Sci. 114, E2466–E2475.

Green, K.M., Linsalata, A.E., Todd, P.K., 2016. RAN translation—What makes it run? Brain Res.

Green KM, Glineburg MR, Kearse MG, Flores BN, Linsalata AE, Fedak SJ, et al. RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response. Nat Commun 2017; 8: 2005.

Grima, J.C., Daigle, J.G., Arbez, N., Cunningham, K.C., Zhang, K., Ochaba, J., Geater, C., Morozko, E., Stocksdale, J., Glatzer, J.C., Pham, J.T., Ahmed, I., Peng, Q., Wadhwa, H., Pletnikova, O., Troncoso, J.C., Duan, W., Snyder, S.H., Ranum, L.P.W., Thompson, L.M., Lloyd, T.E., Ross, C.A., Rothstein, J.D., 2017. Mutant Huntingtin Disrupts the Nuclear Pore Complex. Neuron 94, 93–107.e6.

Guerreiro, R., Bras, J., Hardy, J., 2015. SnapShot: Genetics of ALS and FTD. Cell 160, 798.e1.

Gutierrez, H., Davies, A.M., 2011. Regulation of neural process growth, elaboration and structural plasticity by NF-κB. Trends Neurosci.

Haeusler, A.R., Donnelly, C.J., Periz, G., Simko, E.A.J., Shaw, P.G., Kim, M.-S., Maragakis, N.J., Troncoso, J.C., Pandey, A., Sattler, R., Rothstein, J.D., Wang, J., 2014. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. Nature 507, 195–200.

Haeusler, A.R., Donnelly, C.J., Rothstein, J.D., 2016. The expanding biology of the C9orf72 nucleotide repeat expansion in neurodegenerative disease. Nat. Rev. Neurosci. 17, 383–395.

Haney, M.S., Kramer, N.J., Morgens, D.W., Jovičić, A., Couthouis, J., Li, A., Ousey, J., Ma, R., Bieri, G., Bassik, M.C., Gitler, A.D., 2017. CRISPR-Cas9 Screens In Human Cells And Primary Neurons Identify Modifiers Of C9orf72 Dipeptide Repeat Protein Toxicity. bioRxiv 129254.

Hardy, J., Rogaeva, E., 2014. Motor neuron disease and frontotemporal dementia: Sometimes related, sometimes not. Exp. Neurol.

Harms, M.B., Cady, J., Zaidman, C., Cooper, P., Bali, T., Allred, P., Cruchaga, C., Baughn, M., Libby, R.T., Pestronk, A., Goate, A., Ravits, J., Baloh, R.H., 2013. Lack of C9ORF72 coding mutations supports a gain of function for repeat expansions in amyotrophic lateral sclerosis. Neurobiol. Aging 34, 2234.e13-9.

Harrison, D.E., Strong, R., Sharp, Z.D., Nelson, J.F., Astle, C.M., Flurkey, K., Nadon, N.L., Wilkinson, J.E., Frenkel, K., Carter, C.S., Pahor, M., Javors, M.A., Fernandez, E., Miller, R.A., 2009. Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. Nature 460, 392–5.

Hautbergue, G.M., Castelli, L.M., Ferraiuolo, L., Sanchez-Martinez, A., Cooper-Knock, J., Higginbottom, A., Lin, Y.-H., Bauer, C.S., Dodd, J.E., Myszczynska, M.A., Alam, S.M., Garneret, P., Chandran, J.S., Karyka, E., Stopford, M.J., Smith, E.F., Kirby, J., Meyer, K., Kaspar, B.K., Isaacs, A.M., El-Khamisy, S.F., De Vos, K.J., Ning, K., Azzouz, M., Whitworth, A.J., Shaw, P.J., 2017. SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits. Nat. Commun. 8, 16063.

Heneka, M.T., Kummer, M.P., Latz, E., 2014. Innate immune activation in neurodegenerative disease. Nat. Rev. Immunol. 14, 463–477.

Hensman, D.J., Poulter, M., Beck, J., Hehir, J., Polke, J.M., Campbell, T., Adamson, G., Mudanohwo, E., McColgan, P., Haworth, A., Wild, E.J., Sweeney, M.G., Houlden, H., Mead, S., Tabrizi, S.J., 2014. C9orf72 expansions are the most common genetic cause of Huntington disease phenocopies. Neurology 82, 292–299.

Hirth, F., 2010. Drosophila melanogaster in the Study of Human Neurodegeneration. CNS Neurol. Disord. -Drug Targets 9, 504–523.

Honda, S., Arakawa, S., Nishida, Y., Yamaguchi, H., Ishii, E., Shimizu, S., 2014. Ulk1-mediated Atg5-independent macroautophagy mediates elimination of mitochondria from embryonic reticulocytes. Nat. Commun. 5, ncomms5004.

Hsiung, G.Y.R., Dejesus-Hernandez, M., Feldman, H.H., Sengdy, P., Bouchard-Kerr, P., Dwosh, E., Butler, R., Leung, B., Fok, A., Rutherford, N.J., Baker, M., Rademakers, R., Mackenzie, I.R.A., 2012. Clinical and pathological features of familial frontotemporal dementia caused by C9ORF72 mutation on chromosome 9p. Brain 135, 709–722.

Hübers, A., Marroquin, N., Schmoll, B., Vielhaber, S., Just, M., Mayer, B., H??gel, J., Dorst, J., Mertens, T., Just, W., Aulitzky, A., Wais, V., Ludolph, A.C., Kubisch, C., Weishaupt, J.H., Volk, A.E., 2014. Polymerase chain reaction and Southern blot-based analysis of the C9orf72 hexanucleotide repeat in different motor neuron diseases. Neurobiol. Aging 35, 1214.e1-1214.e6.

Hülsmann, B.B., Labokha, A.A., Görlich, D., 2012. The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. Cell 150, 738–751.

Hutagalung, A.H., Novick, P.J., 2011. Role of Rab GTPases in membrane traffic and cell physiology. Physiol. Rev. 91, 119–49.

Hyman, A.A., Weber, C.A., Jülicher, F., 2014. Liquid-Liquid Phase Separation in Biology. Annu. Rev. Cell Dev. Biol. 30, 39–58.

Igaz, L.M., Kwong, L.K., Lee, E.B., Chen-Plotkin, A., Swanson, E., Unger, T., Malunda, J., Xu, Y., Winton, M.J., Trojanowski, J.Q., Lee, V.M.Y., 2011. Dysregulation of the ALS-associated gene TDP-43 leads to neuronal death and degeneration in mice. J. Clin. Invest. 121, 726–738.

Ikeda, Y., Dalton, J.C., Moseley, M.L., Gardner, K.L., Bird, T.D., Ashizawa, T., Seltzer, W.K., Pandolfo, M., Milunsky, A., Potter, N.T., Shoji, M., Vincent, J.B., Day, J.W., Ranum, L.P.W., 2004. Spinocerebellar ataxia type 8: molecular genetic comparisons and haplotype analysis of 37 families with ataxia. Am. J. Hum. Genet. 75, 3–16.

Ince, P.G., Evans, J., Knopp, M., Forster, G., Hamdalla, H.H.M., Wharton, S.B., Shaw, P.J., 2003. Corticospinal tract degeneration in the progressive muscular atrophy variant of ALS. Neurology 60, 1252–1258.

Ishiguro, A., Kimura, N., Watanabe, Y., Watanabe, S., Ishihama, A., 2016. TDP-43 binds and transports G-quadruplex-containing mRNAs into neurites for local translation. Genes to Cells 21, 466–481.

Ishiguro, T., Sato, N., Ueyama, M., Fujikake, N., Sellier, C., Kanegami, A., Tokuda, E., Zamiri, B., Gall-Duncan, T., Mirceta, M., Furukawa, Y., Yokota, T., Wada, K., Taylor, J.P., Pearson, C.E., Charlet-Berguerand, N., Mizusawa, H., Nagai, Y., Ishikawa, K., 2017. Regulatory Role of RNA Chaperone TDP-43 for RNA Misfolding and Repeat-Associated Translation in SCA31. Neuron 94, 108–124.e7.

Ishiura, H., Takahashi, Y., Mitsui, J., Yoshida, S., Kihira, T., Kokubo, Y., Kuzuhara, S., Ranum, L.P.W., Tamaoki, T., Ichikawa, Y., Date, H., Goto, J., Tsuji, S., 2012. C9ORF72 Repeat Expansion in Amyotrophic Lateral Sclerosis in the Kii Peninsula of Japan. Arch. Neurol. 69, 1154–8.

Jain, A., Vale, R.D., 2017. RNA phase transitions in repeat expansion disorders. Nature 546, 243–247.

Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., Parker, R., 2016. ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell 164, 487–498.

Jiang, J., Zhu, Q., Gendron, T.F., Saberi, S., McAlonis-Downes, M., Seelman, A., Stauffer, J.E., Jafar-nejad, P., Drenner, K., Schulte, D., Chun, S., Sun, S., Ling, S.C., Myers, B., Engelhardt, J., Katz, M., Baughn, M., Platoshyn, O., Marsala, M., Watt, A., Heyser, C.J., Ard, M.C., De Muynck, L., Daughrity, L.M., Swing, D.A., Tessarollo, L., Jung, C.J., Delpoux, A., Utzschneider, D.T., Hedrick, S.M., de Jong, P.J., Edbauer, D., Van Damme, P., Petrucelli, L., Shaw, C.E., Bennett,

C.F., Da Cruz, S., Ravits, J., Rigo, F., Cleveland, D.W., Lagier-Tourenne, C., 2016. Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. Neuron 90, 535–550.

Jiao, B., Tang, B., Liu, X., Yan, X., Zhou, L., Yang, Y., Wang, J., Xia, K., Shen, L., 2014. Identification of C9orf72 repeat expansions in patients with amyotrophic lateral sclerosis and frontotemporal dementia in mainland China. Neurobiol. Aging 35, 936.e19-936.e22.

Jovičić, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N., Yamada, S.B., Paul, J.W., Sun, S., Herdy, J.R., Bieri, G., Kramer, N.J., Gage, F.H., Van Den Bosch, L., Robberecht, W., Gitler, A.D., 2015. Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. Nat. Neurosci. 18, 1226–1229.

Jungreis, I., Lin, M.F., Spokony, R., Chan, C.S., Negre, N., Victorsen, A., White, K.P., Kellis, M., 2011. Evidence of abundant stop codon readthrough in Drosophila and other metazoa. Genome Res. 21, 2096–2113.

Kabashi, E., Agar, J.N., Strong, M.J., Durham, H.D., 2012. Impaired proteasome function in sporadic amyotrophic lateral sclerosis. Amyotroph. Lateral Scler. 13, 367–371.

Kaehler, C., Isensee, J., Nonhoff, U., Terrey, M., Hucho, T., Lehrach, H., Krobitsch, S., 2012. Ataxin-2-Like Is a Regulator of Stress Granules and Processing Bodies. PLoS One 7, e50134.

Kaneb, H.M., Folkmann, A.W., Belzil, V. V., Jao, L.E., Leblond, C.S., Girard, S.L., Daoud, H., Noreau, A., Rochefort, D., Hince, P., Szuto, A., Levert, A., Vidal, S., Andre-Guimont, C., Camu, W., Bouchard, J.P., Dupre, N., Rouleau, G.A., Wente, S.R., Dion, P.A., 2015. Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis. Hum. Mol. Genet. 24, 1363–1373.

Kanekura, K., Yagi, T., Cammack, A.J., Mahadevan, J., Kuroda, M., Harms, M.B., Miller, T.M., Urano, F., 2016. Poly-dipeptides encoded by the C9ORF72 repeats block global protein translation. Hum. Mol. Genet. 25, 1803–1813.

Kansal, K., Mareddy, M., Sloane, K.L., Minc, A.A., Rabins, P. V, McGready, J.B., Onyike, C.U., 2016. Survival in Frontotemporal Dementia Phenotypes: A Meta-Analysis. Dement. Geriatr. Cogn. Disord.

Karam, C., Scelsa, S.N., MacGowan, D.J.L., 2010. The clinical course of progressive bulbar palsy. Amyotroph. Lateral Scler. 11, 364–368.

Kearse, M.G., Green, K.M., Krans, A., Rodriguez, C.M., Linsalata, A.E., Goldstrohm, A.C., Todd, P.K., 2016. CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. Mol. Cell 62, 314–322.

Kelley, J.B., Talley, A.M., Spencer, A., Gioeli, D., Paschal, B.M., Terry, L., Shows, E., Wente, S., Goldfarb, D., Corbett, A., Mason, D., Harreman, M., Adam, S., Harel, A., Forbes, D., Pemberton, L., Paschal, B., Fornerod, M., Ohno, M., Yoshida, M., Mattaj, I., Lange, A., Mills, R., Lange, C., Stewart, M., Devine, S., Corbett, A., Peifer, M., Berg, S., Reynolds, A., Conti, E., Uy, M., Leighton, L., Blobel, G., Kuriyan, J., Fontes, M., Teh, T., Kobe, B., Conti, E., Kuriyan, J., Moroianu, J., Blobel, G., Radu, A., Harreman, M., Hodel, M., Fanara, P., Hodel, A., Corbett, A., Kobe, B., Fanara, P., Hodel, M., Corbett, A., Hodel, A., Sun, C., Yang, W., Tu, L., Musser, S., Kutay, U., Bischoff, F., Kostka, S., Kraft, R., Gorlich, D., Fagerlund, R., Kinnunen, L., Kohler, M., Julkunen, I., Melen, K., Kohler, M., Gorlich, D., Hartmann, E., Franke, J., Kohler, M., Speck, C., Christiansen, M., Bischoff, F., Prehn, S., Haller, H., Gorlich, D., Hartmann, E., Welch, K., Franke, J., Kohler, M., Macara, I., Friedrich, B., Quensel, C., Sommer, T., Hartmann, E., Kohler, M., Holt, J., Ly-Huynh, J., Efthymiadis, A., Hime, G., Loveland, K., Jans, D., Mason, D., Fleming, R., Goldfarb, D., Mason, D., Mathe, E., Fleming, R., Goldfarb, D., Geles, K., Adam, S., Yasuhara, N., Shibazaki, N., Tanaka, S., Nagai, M., Kamikawa, Y., Oe, S., Asally, M., Kamachi, Y., Kondoh, H., Yoneda, Y., Kohler, M., Ansieau, S., Prehn, S., Leutz, A., Haller, H., Hartmann, E., Cutress, M., Whitaker, H., Mills, I., Stewart, M., Neal, D., Tarendeau, F., Boudet, J., Guilligay, D., Mas, P., Bougault, C., Boulo, S., Baudin, F., Ruigrok, R., Daigle, N., Ellenberg, J., Cusak, S., Simorre, J., Hart, D., Fontes, M., Teh, T., Jans, D., Brinkworth, R., Kobe, B., Fontes, M., Teh, T., Toth, G., John, A., Pavo, I., Jans, D., Kobe, B., Tejomurtula, J., Lee, K., Tripurani, S., Smith, G., Yao, J., 2010. Karyopherin alpha7 (KPNA7), a divergent member of the importin alpha family of nuclear import receptors. BMC Cell Biol. 11, 63.

Khan, B.K., Yokoyama, J.S., Takada, L.T., Sha, S.J., Rutherford, N.J., Fong, J.C., Karydas, A.M., Wu, T., Ketelle, R.S., Baker, M.C., Hernandez, M.-D., Coppola, G., Rademakers, R., Lee, S.E., Rosen, H.J., Rabinovici, G.D., Seeley, W.W., Rankin, K.P., Boxer, A.L., Miller, B.L., 2012. Atypical, slowly progressive behavioural variant frontotemporal dementia associated with C9ORF72 hexanucleotide expansion. J. Neurol. Neurosurg. Psychiatry 83, 358–364.

Khosravi, B., Hartmann, H., May, S., Möhl, C., Ederle, H., Michaelsen, M., Schludi, M.H., Dormann, D., Edbauer, D., 2017. Cytoplasmic poly-GA aggregates impair nuclear import of TDP-43 in C9orf72 ALS/FTLD. Hum. Mol. Genet. 26, 790–800.

Kiernan, M.C., Vucic, S., Cheah, B.C., Turner, M.R., Eisen, A., Hardiman, O., Burrell, J.R., Zoing, M.C., 2011. Amyotrophic lateral sclerosis. In: The Lancet. pp. 942–955.

Kim, H.J., Taylor, J.P., 2017. Lost in Transportation: Nucleocytoplasmic Transport Defects in ALS and Other Neurodegenerative Diseases. Neuron 96, 285–297.

Kim, H.-J., Raphael, A.R., LaDow, E.S., McGurk, L., Weber, R.A., Trojanowski, J.Q., Lee, V.M.-Y., Finkbeiner, S., Gitler, A.D., Bonini, N.M., 2013. Therapeutic modulation of eIF2α phosphorylation rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models. Nat. Genet. 46, 152–160.

Kim, S.H., Shanware, N.P., Bowler, M.J., Tibbetts, R.S., 2010. Amyotrophic lateral sclerosis-associated proteins TDP-43 and FUS/TLS function in a common biochemical complex to co-regulate HDAC6 mRNA. J. Biol. Chem. 285, 34097–34105.

Kim, W.-K., Liu, X., Sandner, J., Pasmantier, M., Andrews, J., Rowland, L.P., Mitsumoto, H., 2009. Study of 962 patients indicates progressive muscular atrophy is a form of ALS. Neurology 73, 1686–1692.

Kino, Y., Washizu, C., Kurosawa, M., Oma, Y., Hattori, N., Ishiura, S., Nukina, N., 2015. Nuclear localization of MBNL1: Splicing-mediated autoregulation and repression of repeat-derived aberrant proteins. Hum. Mol. Genet. 24, 740–756.

Kinoshita, Y., Ito, H., Hirano, A., Fujita, K., Wate, R., Nakamura, M., Kaneko, S., Nakano, S., Kusaka, H., 2009. Nuclear contour irregularity and abnormal transporter protein distribution in anterior horn cells in amyotrophic lateral sclerosis. J. Neuropathol. Exp. Neurol. 68, 1184–1192.

Kodiha, M., Tran, D., Qian, C., Morogan, A., Presley, J.F., Brown, C.M., Stochaj, U., 2008. Oxidative stress mislocalizes and retains transport factor importin-α and nucleoporins Nup153 and Nup88 in nuclei where they generate high molecular mass complexes. Biochim. Biophys. Acta - Mol. Cell Res. 1783, 405–418.

Kohler, M., Speck, C., Christiansen, M., Bischoff, F.R., Prehn, S., Haller, H., Gorlich, D., Hartmann, E., 1999. Evidence for distinct substrate specificities of importin alpha family members in nuclear protein import. Mol Cell Biol 19, 7782–7791.

Koppers, M., Blokhuis, A.M., Westeneng, H.J., Terpstra, M.L., Zundel, C.A.C., Vieira De Sá, R., Schellevis, R.D., Waite, A.J., Blake, D.J., Veldink, J.H., Van Den Berg, L.H., Pasterkamp, R.J., 2015. C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. Ann. Neurol. 78, 426–438.

Kramer, N.J., Carlomagno, Y., Zhang, Y.-J., Almeida, S., Cook, C.N., Gendron, T.F., Prudencio, M., Van Blitterswijk, M., Belzil, V., Couthouis, J., Paul, J.W., Goodman, L.D., Daughrity, L., Chew, J., Garrett, A., Pregent, L., Jansen-West, K., Tabassian, L.J., Rademakers, R., Boylan, K., Graff-Radford, N.R., Josephs, K.A., Parisi, J.E., Knopman, D.S., Petersen, R.C., Boeve, B.F., Deng, N., Feng, Y., Cheng, T.-H., Dickson, D.W., Cohen, S.N., Bonini, N.M., Link, C.D., Gao, F.-B., Petrucelli, L., Gitler, A.D., 2016. Spt4 selectively regulates the expression of C9orf72 sense and antisense mutant transcripts. Science (80-.). 353, 708–712.

Kuipers-Upmeijer, J., 2001. Primary lateral sclerosis: clinical, neurophysiological, and magnetic resonance findings. J. Neurol. Neurosurg. Psychiatry 71, 615–620.

Kumar, J.P., 2012. Building an ommatidium one cell at a time. Dev. Dyn.

Kusano, A., Staber, C., Chan, H.Y.E., Ganetzky, B., 2003. Closing the (Ran)GAP on segregation distortion in Drosophila. BioEssays.

Kwon, I., Xiang, S., Kato, M., Wu, L., Theodoropoulos, P., Wang, T., Kim, J., Yun, J., Xie, Y., McKnight, S.L., 2014. Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. Science (80-.). 345, 1139–1145.

Laaksovirta, H., Peuralinna, T., Schymick, J.C., Scholz, S.W., Lai, S.L., Myllykangas, L., Sulkava, R., Jansson, L., Hernandez, D.G., Gibbs, J.R., Nalls, M.A., Heckerman, D., Tienari, P.J., Traynor, B.J., 2010. Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: A genome-wide association study. Lancet Neurol. 9, 978–985.

Lagier-Tourenne, C., Baughn, M., Rigo, F., Sun, S., Liu, P., Li, H.-R., Jiang, J., Watt, A.T., Chun, S., Katz, M., Qiu, J., Sun, Y., Ling, S.-C., Zhu, Q., Polymenidou, M., Drenner, K., Artates, J.W., McAlonis-Downes, M., Markmiller, S., Hutt, K.R., Pizzo, D.P., Cady, J., Harms, M.B., Baloh, R.H., Vandenberg, S.R., Yeo, G.W., Fu, X.-D., Bennett, C.F., Cleveland, D.W., Ravits, J., 2013. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. Proc. Natl. Acad. Sci. 110, E4530–E4539.

Laitman, B.M., Mariani, J.N., Zhang, C., Sawai, S., John, G.R., 2017. Karyopherin alpha proteins regulate oligodendrocyte differentiation. PLoS One 12, e0170477.

Lattante, S., Ciura, S., Rouleau, G.A., Kabashi, E., 2015. Defining the genetic connection linking amyotrophic lateral sclerosis (ALS) with frontotemporal dementia (FTD). Trends Genet.

Lawrence, T., 2009. The nuclear factor NF-kappaB pathway in inflammation. Cold Spring Harb. Perspect. Biol.

Le Forestier, N., Maisonobe, T., Piquard, A., Rivaud, S., Crevier-Buchman, L., Salachas, F., Pradat, P.F., Lacomblez, L., Meininger, V., 2001. Does primary lateral sclerosis exist? A study of 20 patients and a review of the literature. Brain 124, 1989–1999.

Lee, E.B., Lee, V.M.-Y., Trojanowski, J.Q., 2011. Gains or losses: molecular mechanisms of TDP43-mediated neurodegeneration. Nat. Rev. Neurosci. 13, 38–50.

Lee, K.H., Zhang, P., Kim, H.J., Mitrea, D.M., Sarkar, M., Freibaum, B.D., Cika, J., Coughlin, M., Messing, J., Molliex, A., Maxwell, B.A., Kim, N.C., Temirov, J., Moore, J., Kolaitis, R.M., Shaw, T.I., Bai, B., Peng, J., Kriwacki, R.W., Taylor, J.P., 2016. C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. Cell 167, 774–788.e17.

Lee, S., Huang, E.J., 2017. Modeling ALS and FTD with iPSC-derived neurons. Brain Res.

Lee, Y.B., Chen, H.J., Peres, J.N., Gomez-Deza, J., Attig, J., Stalekar, M., Troakes, C., Nishimura, A.L., Scotter, E.L., Vance, C., Adachi, Y., Sardone, V., Miller,

J.W., Smith, B.N., Gallo, J.M., Ule, J., Hirth, F., Rogelj, B., Houart, C., Shaw, C.E., 2013. Hexanucleotide repeats in ALS/FTD form length-dependent RNA Foci, sequester RNA binding proteins, and are neurotoxic. Cell Rep. 5, 1178–1186.

Lee, Y.-B., Baskaran, P., Gomez, J., Chen, H.-J., Nishimura, A., Smith, B., Troakes, C., Adachi, Y., Stepto, A., Petrucelli, L., Gallo, J.-M., Hirth, F., Rogelj, B., Guthrie, S., Shaw, C.E., 2017. C9orf72 poly GA RAN-translated protein plays a key role in Amyotrophic Lateral Sclerosis via aggregation and toxicity. Hum. Mol. Genet.

Lehmer, C., Oeckl, P., Weishaupt, J.H., Volk, A.E., Diehl-Schmid, J., Schroeter, M.L., Lauer, M., Kornhuber, J., Levin, J., Fassbender, K., Landwehrmeyer, B., Schludi, M.H., Arzberger, T., Kremmer, E., Flatley, A., Feederle, R., Steinacker, P., Weydt, P., Ludolph, A.C., Edbauer, D., Otto, M., 2017. Poly-GP in cerebrospinal fluid links C9orf72 -associated dipeptide repeat expression to the asymptomatic phase of ALS/FTD. EMBO Mol. Med. 9, 859–868.

Lesage, S., Le Ber, I., Condroyer, C., Broussolle, E., Gabelle, A., Thobois, S., Pasquier, F., Mondon, K., Dion, P.A., Rochefort, D., Rouleau, G.A., D??rr, A., Brice, A., 2013. C9orf72 repeat expansions are a rare genetic cause of parkinsonism. Brain 136, 385–391.

Levine, T.P., Daniels, R.D., Gatta, A.T., Wong, L.H., Hayes, M.J., 2013. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. Bioinformatics 29, 499–503.

Li, L.-B., Yu, Z., Teng, X., Bonini, N.M., 2008. RNA toxicity is a component of ataxin-3 degeneration in Drosophila. Nature 453, 1107–1111.

Li, Y.R., King, O.D., Shorter, J., Gitler, A.D., 2013. Stress granules as crucibles of ALS pathogenesis. J. Cell Biol.

Lillo, P., Hodges, J.R., 2009. Frontotemporal dementia and motor neurone disease: Overlapping clinic-pathological disorders. J. Clin. Neurosci.

Lin, C.L.G., Bristol, L.A., Jin, L., Dykes-Hoberg, M., Crawford, T., Clawson, L., Rothstein, J.D., 1998. Aberrant RNA processing in a neurodegenerative disease: The cause for absent EAAT2, a glutamate transporter, in amyotrophic lateral sclerosis. Neuron 20, 589–602.

Lin, Y., Mori, E., Kato, M., Xiang, S., Wu, L., Kwon, I., McKnight, S.L., 2016. Toxic PR Poly-Dipeptides Encoded by the C9orf72 Repeat Expansion Target LC Domain Polymers. Cell 167, 789–802.e12.

Lin, Y., Protter, D.S.W., Rosen, M.K., Parker, R., 2015. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol. Cell 60, 208–219.

Lindquist, S.G., Duno, M., Batbayli, M., Puschmann, A., Braendgaard, H., Mardosiene, S., Svenstrup, K., Pinborg, L.H., Vestergaard, K., Hjermind, L.E., Stokholm, J., Andersen, B.B., Johannsen, P., Nielsen, J.E., 2013. Corticobasal and ataxia syndromes widen the spectrum of C9ORF72 hexanucleotide expansion disease. Clin. Genet. 83, 279–283.

Ling, S.C., Polymenidou, M., Cleveland, D.W., 2013. Converging mechanisms in ALS and FTD: Disrupted RNA and protein homeostasis. Neuron 79, 416–438.

Lippai, M., Tirian, L., Boros, I., Mihaly, J., Erdelyi, M., Belecz, I., Mathe, E., Posfai, J., Nagy, A., Udvardy, A., Paraskeva, E., Gorlich, D., Szabad, J., 2000. The Ketel gene encodes a Drosophila homologue of importin-beta. Genetics 156, 1889–1900.

Liu, E.Y., Russ, J., Wu, K., Neal, D., Suh, E., McNally, A.G., Irwin, D.J., Van Deerlin, V.M., Lee, E.B., 2014. C9orf72 hypermethylation protects against repeat expansion-associated pathology in ALS/FTD. Acta Neuropathol. 128, 525–541.

Liu, F., Liu, Q., Lu, C.X., Cui, B., Guo, X.N., Wang, R.R., Liu, M.S., Li, X.G., Cui, L. ying, Zhang, X., 2016a. Identification of a novel loss-of-function C9orf72 splice site mutation in a patient with amyotrophic lateral sclerosis. Neurobiol. Aging 47, 219.e1-219.e5.

Liu, Q., Shu, S., Wang, R.R., Liu, F., Cui, B., Guo, X.N., Lu, C.X., Li, X.G., Liu, M.S., Peng, B., Cui, L.Y., Zhang, X., 2016b. Whole-exome sequencing identifies a missense mutation in hnRNPA1 in a family with flail arm ALS. Neurology 87, 1763–1769.

Liu, Y., Pattamatta, A., Zu, T., Reid, T., Bardhi, O., Borchelt, D.R., Yachnis, A.T., Ranum, L.P.W., 2016c. C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD. Neuron 90, 521–534.

Liu-Yesucevitz, L., Bilgutay, A., Zhang, Y.J., Vanderwyde, T., Citro, A., Mehta, T., Zaarur, N., McKee, A., Bowser, R., Sherman, M., Petrucelli, L., Wolozin, B., 2010. Tar DNA binding protein-43 (TDP-43) associates with stress granules: Analysis of cultured cells and pathological brain tissue. PLoS One 5, e13250.

Lo Giudice, T., Lombardi, F., Santorelli, F.M., Kawarai, T., Orlacchio, A., 2014. Hereditary spastic paraplegia: Clinical-genetic characteristics and evolving molecular mechanisms. Exp. Neurol.

Logroscino, G., Traynor, B.J., Hardiman, O., Chio, A., Mitchell, D., Swingler, R.J., Millul, A., Benn, E., Beghi, E., 2010. Incidence of amyotrophic lateral sclerosis in Europe. J. Neurol. Neurosurg. Psychiatry 81, 385–390.

Lopez-Gonzalez, R., Lu, Y., Gendron, T.F., Karydas, A., Tran, H., Yang, D., Petrucelli, L., Miller, B.L., Almeida, S., Gao, F.B., 2016. Poly(GR) in C9ORF72-Related ALS/FTD Compromises Mitochondrial Function and Increases Oxidative Stress and DNA Damage in iPSC-Derived Motor Neurons. Neuron 92, 383–391.

Maas, S., Gommans, W.M., 2009. Identification of a selective nuclear import signal in adenosine deaminases acting on RNA. Nucleic Acids Res. 37, 5822–5829.

Mackenzie, I.R.A., Frick, P., Grässer, F.A., Gendron, T.F., Petrucelli, L., Cashman, N.R., Edbauer, D., Kremmer, E., Prudlo, J., Troost, D., Neumann, M., 2015. Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. Acta Neuropathol. 130, 845–861.

Mackenzie, I.R.A., Frick, P., Neumann, M., 2014. The neuropathology associated with repeat expansions in the C9ORF72 gene. Acta Neuropathol.

MacKenzie, I.R., Arzberger, T., Kremmer, E., Troost, D., Lorenzl, S., Mori, K., Weng, S.M., Haass, C., Kretzschmar, H.A., Edbauer, D., Neumann, M., 2013. Dipeptide repeat protein pathology in C9ORF72 mutation cases: Clinico-pathological correlations. Acta Neuropathol. 126, 859–879.

MacKenzie, I.R., Arzberger, T., Kremmer, E., Troost, D., Lorenzl, S., Mori, K., Weng, S.M., Haass, C., Kretzschmar, H.A., Edbauer, D., Neumann, M., 2013. Dipeptide repeat protein pathology in C9ORF72 mutation cases: Clinico-pathological correlations. Acta Neuropathol. 126, 859–879.

Mackenzie, I.R., Nicholson, A.M., Sarkar, M., Messing, J., Purice, M.D., Pottier, C., Annu, K., Baker, M., Perkerson, R.B., Kurti, A., Matchett, B.J., Mittag, T., Temirov, J., Hsiung, G.Y.R., Krieger, C., Murray, M.E., Kato, M., Fryer, J.D., Petrucelli, L., Zinman, L., Weintraub, S., Mesulam, M., Keith, J., Zivkovic, S.A., Hirsch-Reinshagen, V., Roos, R.P., Züchner, S., Graff-Radford, N.R., Petersen, R.C., Caselli, R.J., Wszolek, Z.K., Finger, E., Lippa, C., Lacomis, D., Stewart, H., Dickson, D.W., Kim, H.J., Rogaeva, E., Bigio, E., Boylan, K.B., Taylor, J.P., Rademakers, R., 2017. TIA1 Mutations in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase Separation and Alter Stress Granule Dynamics. Neuron 95, 808–816.e9.

Mackenzie, I.R.A., Neumann, M., 2016. Molecular neuropathology of frontotemporal dementia: insights into disease mechanisms from postmortem studies. J. Neurochem.

Macnair, L., Xiao, S., Miletic, D., Ghani, M., Julien, J.P., Keith, J., Zinman, L., Rogaeva, E., Robertson, J., 2016. MTHFSD and DDX58 are novel RNA-binding proteins abnormally regulated in amyotrophic lateral sclerosis. Brain 139, 86–100.

Maharjan, N., Künzli, C., Buthey, K., Saxena, S., 2017. C9ORF72 Regulates Stress Granule Formation and Its Deficiency Impairs Stress Granule Assembly, Hypersensitizing Cells to Stress. Mol. Neurobiol. 54, 3062–3077.

Mahboubi, H., Seganathy, E., Kong, D., Stochaj, U., 2013. Identification of Novel Stress Granule Components That Are Involved in Nuclear Transport. PLoS One 8, e68356.

Mahoney, C.J., Beck, J., Rohrer, J.D., Lashley, T., Mok, K., Shakespeare, T., Yeatman, T., Warrington, E.K., Schott, J.M., Fox, N.C., Rossor, M.N., Hardy, J., Collinge, J., Revesz, T., Mead, S., Warren, J.D., 2012. Frontotemporal dementia with the C9ORF72 hexanucleotide repeat expansion: Clinical, neuroanatomical and neuropathological features. Brain 135, 736–750.

Majounie, E., Abramzon, Y., Renton, A.E., Keller, M.F., Traynor, B.J., Singleton, A.B., 2012a. Large C9orf72 repeat expansions are not a common cause of Parkinson's disease. Neurobiol. Aging 33, 2527.e1-2.

Majounie, E., Renton, A.E., Mok, K., Dopper, E.G., Waite, A., Rollinson, S., Chiò, A., Restagno, G., Nicolaou, N., Simon-Sanchez, J., van Swieten, J.C., Abramzon, Y., Johnson, J.O., Sendtner, M., Pamphlett, R., Orrell, R.W., Mead, S., Sidle, K.C., Houlden, H., Rohrer, J.D., Morrison, K.E., Pall, H., Talbot, K., Ansorge, O., Hernandez, D.G., Arepalli, S., Sabatelli, M., Mora, G., Corbo, M., Giannini, F., Calvo, A., Englund, E., Borghero, G., Floris, G.L., Remes, A.M., Laaksovirta, H., McCluskey, L., Trojanowski, J.Q., Van Deerlin, V.M., Schellenberg, G.D., Nalls, M.A., Drory, V.E., Lu, C.S., Yeh, T.H., Ishiura, H., Takahashi, Y., Tsuji, S., Le Ber, I., Brice, A., Drepper, C., Williams, N., Kirby, J., Shaw, P., Hardy, J., Tienari, P.J., Heutink, P., Morris, H.R., Pickering-Brown, S., Traynor, B.J., 2012b. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: A cross-sectional study. Lancet Neurol. 11, 323–330.

Majumder, P., Chen, Y.T., Bose, J.K., Wu, C.C., Cheng, W.C., Cheng, S.J., Fang, Y.H., Chen, Y.L., Tsai, K.J., Lien, C.C., Shen, C.K.J., 2012. TDP-43 regulates the mammalian spinogenesis through translational repression of Rac1. Acta Neuropathol. 124, 231–245.

Majumder, P., Chu, J.F., Chatterjee, B., Swamy, K.B.S., Shen, C.K.J., 2016. Co-regulation of mRNA translation by TDP-43 and Fragile X Syndrome protein FMRP. Acta Neuropathol. 132, 721–738.

Mann, D.M., Rollinson, S., Robinson, A., Bennion Callister, J., Thompson, J.C., Snowden, J.S., Gendron, T., Petrucelli, L., Masuda-Suzukake, M., Hasegawa, M., Davidson, Y., Pickering-Brown, S., 2013. Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. Acta Neuropathol. Commun. 1, 68.

Markmiller, S., Soltanieh, S., Server, L.K., Mak, R., Wenhao, J., Fang, Y.M., Luo, E., Krach, F., Yang, D., Anindya, S., Fulzele, A., Wozniak, M.J., Gonzalez, J.D., Kankel, W.M., Gao, B.F., Bennett, J.E., Lécuyer, E., Yeo, W.Y., 2018. Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. Cell 172, 590–604.

Maruyama, H., Morino, H., Ito, H., Izumi, Y., Kato, H., Watanabe, Y., Kinoshita, Y., Kamada, M., Nodera, H., Suzuki, H., Komure, O., Matsuura, S., Kobatake, K., Morimoto, N., Abe, K., Suzuki, N., Aoki, M., Kawata, A., Hirai, T., Kato, T., Ogasawara, K., Hirano, A., Takumi, T., Kusaka, H., Hagiwara, K., Kaji, R.,

Kawakami, H., 2010. Mutations of optineurin in amyotrophic lateral sclerosis. Nature 465, 223–226.

Matoulkova, E., Michalova, E., Vojtesek, B., Hrstka, R., 2012. The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. RNA Biol. 9, 563–576.

May, S., Hornburg, D., Schludi, M.H., Arzberger, T., Rentzsch, K., Schwenk, B.M., Grässer, F.A., Mori, K., Kremmer, E., Banzhaf-Strathmann, J., Mann, M., Meissner, F., Edbauer, D., 2014. C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. Acta Neuropathol. 128, 485–503.

McGurk, L., Lee, V.M., Trojanowksi, J.Q., Van Deerlin, V.M., Lee, E.B., Bonini, N.M., Deerlin, V.M. Van, Lee, E.B., Bonini, N.M., 2014. Poly-A Binding Protein-1 Localization to a Subset of TDP-43 Inclusions in Amyotrophic Lateral Sclerosis Occurs More Frequently in Patients Harboring an Expansion in C9orf72. J. Neuropathol. Exp. Neurol. 73, 1–9.

McInerney, G.M., Kedersha, N.L., Kaufman, R.J., Anderson, P., Liljeström, P., 2005. Importance of eIF2alpha phosphorylation and stress granule assembly in alphavirus translation regulation. Mol. Biol. Cell 16, 3753–63.

McMillan, C.T., Russ, J., Wood, E.M., Irwin, D.J., Grossman, M., McCluskey, L., Elman, L., Van Deerlin, V., Lee, E.B., 2015. C9orf72 promoter hypermethylation is neuroprotective. Neurology 84, 1622–1630.

Meeter, L.H., Kaat, L.D., Rohrer, J.D., van Swieten, J.C., 2017. Imaging and fluid biomarkers in frontotemporal dementia. Nat. Rev. Neurol. 13, 406–419.

Meisler, M.H., Grant, A.E., Jones, J.M., Lenk, G.M., He, F., Todd, P.K., Kamali, M., Albin, R.L., Lieberman, A.P., Langenecker, S.A., Mcinnis, M.G., 2013. C9ORF72 expansion in a family with bipolar disorder. Bipolar Disord. 15, 326–332.

Mertens, J., Paquola, A.C.M., Ku, M., Hatch, E., Böhnke, L., Ladjevardi, S., McGrath, S., Campbell, B., Lee, H., Herdy, J.R., Gonçalves, J.T., Toda, T., Kim, Y., Winkler, J., Yao, J., Hetzer, M.W., Gage, F.H., 2015. Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Reveal Age-Related Nucleocytoplasmic Defects. Cell Stem Cell 17, 705–718.

Mihalas, B.P., Western, P.S., Loveland, K.L., McLaughlin, E.A., Holt, J.E., 2015. Changing expression and subcellular distribution of karyopherins during murine oogenesis. Reproduction 150, 485–496.

Millecamps, S., Boillée, S., Le Ber, I., Seilhean, D., Teyssou, E., Giraudeau, M., Moigneu, C., Vandenberghe, N., Danel-Brunaud, V., Corcia, P., Pradat, P.-F., Le Forestier, N., Lacomblez, L., Bruneteau, G., Camu, W., Brice, A., Cazeneuve, C., LeGuern, E., Meininger, V., Salachas, F., 2012. Phenotype

difference between ALS patients with expanded repeats in C9ORF72 and patients with mutations in other ALS-related genes. J. Med. Genet. 49, 258–263.

Min, J.-N., Whaley, R.A., Sharpless, N.E., Lockyer, P., Portbury, A.L., Patterson, C., 2008. CHIP deficiency decreases longevity, with accelerated aging phenotypes accompanied by altered protein quality control. Mol. Cell. Biol. 28, 4018–25.

Mizielinska, S., Gronke, S., Niccoli, T., Ridler, C.E., Clayton, E.L., Devoy, A., Moens, T., Norona, F.E., Woollacott, I.O.C., Pietrzyk, J., Cleverley, K., Nicoll, A.J., Pickering-Brown, S., Dols, J., Cabecinha, M., Hendrich, O., Fratta, P., Fisher, E.M.C., Partridge, L., Isaacs, A.M., 2014. C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. Science (80-.). 345, 1192–1194.

Mizielinska, S., Lashley, T., Norona, F.E., Clayton, E.L., Ridler, C.E., Fratta, P., Isaacs, A.M., 2013. C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. Acta Neuropathol. 126, 845–857.

Mizielinska, S., Ridler, C.E., Balendra, R., Thoeng, A., Woodling, N.S., Grässer, F.A., Plagnol, V., Lashley, T., Partridge, L., Isaacs, A.M., 2017. Bidirectional nucleolar dysfunction in C9orf72 frontotemporal lobar degeneration. Acta Neuropathol. Commun. 5, 29.

Moens, T.G., Mizielinska, S., Niccoli, T., Mitchell, J.S., Thoeng, A., Ridler, C.E., Grönke, S., Esser, J., Heslegrave, A., Zetterberg, H., Partridge, L., Isaacs, A.M., 2018. Sense and antisense RNA are not toxic in Drosophila models of C9orf72-associated ALS/FTD. Acta Neuropathol. 135, 445–457.

Moens, T.G., Partridge, L., Isaacs, A.M., 2017. Genetic models of C9orf72: what is toxic? Curr. Opin. Genet. Dev.

Mollereau, B., Dominguez, M., Webel, R., Colley, J.N., Keung, B., de Celis, F.J., Desplan, C., 2001. Two-step process for photoreceptor formation in Drosophila. Nature 412, 911–913.

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., Taylor, J.P., 2015. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. Cell 163, 123–133.

Montuschi, A., Iazzolino, B., Calvo, A., Moglia, C., Lopiano, L., Restagno, G., Brunetti, M., Ossola, I., Lo Presti, A., Cammarosano, S., Canosa, A., Chio, A., 2015. Cognitive correlates in amyotrophic lateral sclerosis: a population-based study in Italy. J. Neurol. Neurosurg. Psychiatry 86, 168–173.

Mori, K., Weng, S.-M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., Schmid, B., Kretzschmar, H.A., Cruts, M., Van Broeckhoven, C., Haass, C., Edbauer, D., 2013a. The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS. Science (80-.). 339, 1335–1338.

Mori, K., Arzberger, T., Grässer, F.A., Gijselinck, I., May, S., Rentzsch, K., Weng, S.M., Schludi, M.H., Van Der Zee, J., Cruts, M., Van Broeckhoven, C., Kremmer, E., Kretzschmar, H.A., Haass, C., Edbauer, D., 2013b. Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. Acta Neuropathol. 126, 881–893.

Mori, K., Nihei, Y., Arzberger, T., Zhou, Q., Mackenzie, I.R., Hermann, A., Hanisch, F., Kamp, F., Nuscher, B., Orozco, D., Edbauer, D., Haass, C., 2016. Reduced hnRNPA3 increases C9orf72 repeat RNA levels and dipeptide-repeat protein deposition. EMBO Rep. 17, 1314–1325.

Morimoto, R.I., Cuervo, A.M., 2014. Proteostasis and the aging proteome in health and disease. Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.

Morita, M., Al-Chalabi, A., Andersen, P.M., Hosler, B., Sapp, P., Englund, E., Mitchell, J.E., Habgood, J.J., De Belleroche, J., Xi, J., Jongjaroenprasert, W., Horvitz, H.R., Gunnarsson, L.-G., Brown, R.H., 2006. A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. Neurology 66, 839–844.

Morley, J.F., 2003. Regulation of Longevity in Caenorhabditis elegans by Heat Shock Factor and Molecular Chaperones. Mol. Biol. Cell 15, 657–664.

Morrow, G., Samson, M., Michaud, S., Tanguay, R.M., 2004. Overexpression of the small mitochondrial Hsp22 extends Drosophila life span and increases resistance to oxidative stress. FASEB J. 18, 598–599.

Murakami, T., Qamar, S., Lin, J.Q., Schierle, G.S.K., Rees, E., Miyashita, A., Costa, A.R., Dodd, R.B., Chan, F.T.S., Michel, C.H., Kronenberg-Versteeg, D., Li, Y., Yang, S.P., Wakutani, Y., Meadows, W., Ferry, R.R., Dong, L., Tartaglia, G.G., Favrin, G., Lin, W.L., Dickson, D.W., Zhen, M., Ron, D., Schmitt-Ulms, G., Fraser, P.E., Shneider, N.A., Holt, C., Vendruscolo, M., Kaminski, C.F., St George-Hyslop, P., 2015. ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. Neuron 88, 678–690.

Murphy, N.A., Arthur, K.C., Tienari, P.J., Houlden, H., Chiò, A., Traynor, B.J., 2017. Age-related penetrance of the C9orf72 repeat expansion. Sci. Rep. 7, 2116.

Murray, M.E., Dejesus-Hernandez, M., Rutherford, N.J., Baker, M., Duara, R., Graff-Radford, N.R., Wszolek, Z.K., Ferman, T.J., Josephs, K.A., Boylan, K.B., Rademakers, R., Dickson, D.W., 2011. Clinical and neuropathologic heterogeneity of c9FTD/ALS associated with hexanucleotide repeat expansion in C9ORF72. Acta Neuropathol. 122, 673–690.

Nassif, M., Woehlbier, U., Manque, P.A., 2017. The enigmatic role of C9ORF72 in autophagy. Front. Neurosci.

Nedjic, J., Aichinger, M., Emmerich, J., Mizushima, N., Klein, L., 2008. Autophagy in thymic epithelium shapes the T-cell repertoire and is essential for tolerance. Nature 455, 396–400.

Neumann, M., Sampathu, D.M., Kwong, L.K., Truax, A.C., Micsenyi, M.C., Chou, T.T., Bruce, J., Schuck, T., Grossman, M., Clark, C.M., McCluskey, L.F., Miller, B.L., Masliah, E., Mackenzie, I.R., Feldman, H., Feiden, W., Kretzschmar, H.A., Trojanowski, J.Q., Lee, V.M.-Y., 2006. Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. Science (80-.). 314, 130–133.

Neumann, M., Valori, C.F., Ansorge, O., Kretzschmar, H.A., Munoz, D.G., Kusaka, H., Yokota, O., Ishihara, K., Ang, L.C., Bilbao, J.M., MacKenzie, I.R.A., 2012. Transportin 1 accumulates specifically with FET proteins but no other transportin cargos in FTLD-FUS and is absent in FUS inclusions in ALS with FUS mutations. Acta Neuropathol. 124, 705–716.

Nezis, I.P., Simonsen, A., Sagona, A.P., Finley, K., Gaumer, S., Contamine, D., Rusten, T.E., Stenmark, H., Brech, A., 2008. Ref(2)P, the Drosophila melanogaster homologue of mammalian p62, is required for the formation of protein aggregates in adult brain. J. Cell Biol. 180, 1065–1071.

Niblock, M., Smith, B.N., Lee, Y.-B., Sardone, V., Topp, S., Troakes, C., Al-Sarraj, S., Leblond, C.S., Dion, P.A., Rouleau, G.A., Shaw, C.E., Gallo, J.-M., 2016. Retention of hexanucleotide repeat-containing intron in C9orf72 mRNA: implications for the pathogenesis of ALS/FTD. Acta Neuropathol. Commun. 4, 18.

Nishimura, A.L., Upunski, V., Troakes, C., Kathe, C., Fratta, P., Howell, M., Gallo, J.M., Hortobágyi, T., Shaw, C.E., Rogelj, B., 2010. Nuclear import impairment causes cytoplasmic trans-activation response DNA-binding protein accumulation and is associated with frontotemporal lobar degeneration. Brain 133, 1763–1771.

Niu, C., Zhang, J., Gao, F., Yang, L., Jia, M., Zhu, H., Gong, W., 2012. FUS-NLS/Transportin 1 Complex Structure Provides Insights into the Nuclear Targeting Mechanism of FUS and the Implications in ALS. PLoS One 7, e47056.

Nordin, A., Akimoto, C., Wuolikainen, A., Alstermark, H., Jonsson, P., Birve, A., Marklund, S.L., Graffmo, K.S., Forsberg, K., Brännström, T., Andersen, P.M., 2014. Extensive size variability of the GGGGCC expansion in C9orf72 in both neuronal and non-neuronal tissues in 18 patients with ALS or FTD. Hum. Mol. Genet. 24, 3133–3142.

Nousiainen, H.O., Kestilä, M., Pakkasjärvi, N., Honkala, H., Kuure, S., Tallila, J., Vuopala, K., Ignatius, J., Herva, R., Peltonen, L., 2008. Mutations in mRNA export mediator GLE1 result in a fetal motoneuron disease. Nat. Genet. 40, 155–157.

Oakes, J.A., Davies, M.C., Collins, M.O., 2017. TBK1: a new player in ALS linking autophagy and neuroinflammation. Mol. Brain 10, 5.

Ofer, N., Weisman-Shomer, P., Shklover, J., Fry, M., 2009. The quadruplex r(CGG)n destabilizing cationic porphyrin TMPyP4 cooperates with hnRNPs to increase the translation efficiency of fragile X premutation mRNA. Nucleic Acids Res. 37, 2712–2722.

Ogaki, K., Li, Y., Atsuta, N., Tomiyama, H., Funayama, M., Watanabe, H., Nakamura, R., Yoshino, H., Yato, S., Tamura, A., Naito, Y., Taniguchi, A., Fujita, K., Izumi, Y., Kaji, R., Hattori, N., Sobue, G., 2012. Analysis of C9orf72 repeat expansion in 563 Japanese patients with amyotrophic lateral sclerosis. Neurobiol. Aging 33, 2527.e11-2527.e16.

Ohki, Y., Wenninger-Weinzierl, A., Hruscha, A., Asakawa, K., Kawakami, K., Haass, C., Edbauer, D., Schmid, B., 2017. Glycine-alanine dipeptide repeat protein contributes to toxicity in a zebrafish model of C9orf72 associated neurodegeneration. Mol. Neurodegener. 12, 6.

Ohta, Y., Tremblay, C., Schneider, J.A., Bennett, D.A., Calon, F., Julien, J.-P., 2014. Interaction of transactive response DNA binding protein 43 with nuclear factor κB in mild cognitive impairment with episodic memory deficits. Acta Neuropathol. Commun. 2, 37.

Olney, R.K., Murphy, J., Forshew, D., Garwood, E., Miller, B.L., Langmore, S., Kohn, M.A., Lomen-Hoerth, C., 2005. The effects of executive and behavioral dysfunction on the course of ALS. Neurology 65, 1774–1777.

ORourke, J.G., Bogdanik, L., Yanez, A., Lall, D., Wolf, A.J., Muhammad, A.K.M.G., Ho, R., Carmona, S., Vit, J.P., Zarrow, J., Kim, K.J., Bell, S., Harms, M.B., Miller, T.M., Dangler, C.A., Underhill, D.M., Goodridge, H.S., Lutz, C.M., Baloh, R.H., 2016. C9orf72 is required for proper macrophage and microglial function in mice. Science (80-.). 351, 1324–1329.

O'Rourke, J.G., Bogdanik, L., Muhammad, A.K.M.G., Gendron, T.F., Kim, K.J., Austin, A., Cady, J., Liu, E.Y., Zarrow, J., Grant, S., Ho, R., Bell, S., Carmona, S., Simpkinson, M., Lall, D., Wu, K., Daughrity, L., Dickson, D.W., Harms, M.B., Petrucelli, L., Lee, E.B., Lutz, C.M., Baloh, R.H., 2015. C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. Neuron 88, 892–901.

Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M., Pozniakovski, A., Poser, I., Maghelli, N., Royer, L.A., Weigert, M., Myers, E.W., Grill, S., Drechsel, D., Hyman, A.A., Alberti, S., 2015. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell 162, 1066–1077.

Pearson, J.P., Williams, N.M., Majounie, E., Waite, A., Stott, J., Newsway, V., Murray, A., Hernandez, D., Guerreiro, R., Singleton, A.B., Neal, J., Morris, H.R., 2011. Familial frontotemporal dementia with amyotrophic lateral sclerosis and a shared haplotype on chromosome 9p. J. Neurol. 258, 647–655.

Pelechano, V., Steinmetz, L.M., 2013. Gene regulation by antisense transcription. Nat. Rev. Genet. 14, 880–893.

Perry, D.C., Miller, B.L., 2013. Frontotemporal dementia. Semin Neurol 33, 336–341.

Peters, O.M., Cabrera, G.T., Tran, H., Gendron, T.F., McKeon, J.E., Metterville, J., Weiss, A., Wightman, N., Salameh, J., Kim, J., Sun, H., Boylan, K.B., Dickson, D., Kennedy, Z., Lin, Z., Zhang, Y.J., Daughrity, L., Jung, C., Gao, F.B., Sapp, P.C., Horvitz, H.R., Bosco, D.A., Brown, S.P., de Jong, P., Petrucelli, L., Mueller, C., Brown, R.H., 2015. Human C9ORF72 Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice. Neuron 88, 902–909.

Pettersson, O.J., Aagaard, L., Jensen, T.G., Damgaard, C.K., 2015. Molecular mechanisms in DM1 - A focus on foci. Nucleic Acids Res.

Phukan, J., Pender, N.P., Hardiman, O., 2007. Cognitive impairment in amyotrophic lateral sclerosis. Lancet Neurol.

Pliner, H.A., Mann, D.M., Traynor, B.J., 2014. Searching for Grendel: Origin and global spread of the C9ORF72 repeat expansion. Acta Neuropathol. 127, 391–396.

Pringle, C.E., Hudson, A.J., Munoz, D.G., Kiernan, J.A., Brown, W.F., Ebers, G.C., 1992. Primary lateral sclerosis: Clinical features, neuropathology and diagnostic criteria. Brain 115, 495–520.

Prior, T.W., Snyder, P.J., Rink, B.D., Pearl, D.K., Pyatt, R.E., Mihal, D.C., Conlan, T., Schmalz, B., Montgomery, L., Ziegler, K., Noonan, C., Hashimoto, S., Garner, S., 2010. Newborn and carrier screening for spinal muscular atrophy. Am. J. Med. Genet. Part A 152, 1608–1616.

Protter, D.S.W., Parker, R., 2016. Principles and Properties of Stress Granules. Trends Cell Biol. 26, 668-679.

Proudfoot, M., Gutowski, N.J., Edbauer, D., Hilton, D.A., Stephens, M., Rankin, J., Mackenzie, I.R.A., 2014. Early dipeptide repeat pathology in a frontotemporal dementia kindred with C9ORF72 mutation and intellectual disability. Acta Neuropathol. 127, 451–458.

Prpar Mihevc, S., Darovic, S., Kovanda, A., Bajc Česnik, A., Župunski, V., Rogelj, B., 2017. Nuclear trafficking in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. Brain. 140, 13-26.

Prudencio, M., Belzil, V. V, Batra, R., Ross, C.A., Gendron, T.F., Pregent, L.J., Murray, M.E., Overstreet, K.K., Piazza-Johnston, A.E., Desaro, P., Bieniek, K.F., DeTure, M., Lee, W.C., Biendarra, S.M., Davis, M.D., Baker, M.C., Perkerson, R.B., van Blitterswijk, M., Stetler, C.T., Rademakers, R., Link, C.D., Dickson, D.W., Boylan, K.B., Li, H., Petrucelli, L., 2015. Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. Nat. Neurosci. 18, 1175–1182.

Pujol, G., Soderqvist, H., Radu, A., 2002. Age-associated reduction of nuclear protein import in human fibroblasts. Biochem. Biophys. Res. Commun. 294, 354–358.

Pumroy, R.A., Cingolani, G., 2015. Diversification of importin-α isoforms in cellular trafficking and disease states. Biochem. J 466, 13–28.

Pupillo, E., Messina, P., Logroscino, G., Beghi, E., 2014. Long-term survival in amyotrophic lateral sclerosis: A population-based study. Ann. Neurol. 75, 287–297.

Rabinovici, G.D., Miller, B.L., 2010. Frontotemporal lobar degeneration: Epidemiology, pathophysiology, diagnosis and management. CNS Drugs.

Ramaswami, M., Taylor, J.P., Parker, R., 2013. Altered ribostasis: RNA-protein granules in degenerative disorders. Cell.

Ratti, A., Corrado, L., Castellotti, B., Del Bo, R., Fogh, I., Cereda, C., Tiloca, C., D'Ascenzo, C., Bagarotti, A., Pensato, V., Ranieri, M., Gagliardi, S., Calini, D., Mazzini, L., Taroni, F., Corti, S., Ceroni, M., Oggioni, G.D., Lin, K., Powell, J.F., Sorarù, G., Ticozzi, N., Comi, G.P., D'Alfonso, S., Gellera, C., Silani, V., 2012. C9ORF72 repeat expansion in a large Italian ALS cohort: Evidence of a founder effect. Neurobiol. Aging 33, 2528.e7-2528.e14.

Reddy, K., Zamiri, B., Stanley, S.Y.R., Macgregor, R.B., Pearson, C.E., 2013. The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. J. Biol. Chem. 288, 9860–9866.

Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Laaksovirta, H., van Swieten, J.C., Myllykangas, L., Kalimo, H., Paetau, A., Abramzon, Y., Remes, A.M., Kaganovich, A., Scholz, S.W., Duckworth, J., Ding, J., Harmer, D.W., Hernandez, D.G., Johnson, J.O., Mok, K., Ryten, M., Trabzuni, D., Guerreiro, R.J., Orrell, R.W., Neal, J., Murray, A., Pearson, J., Jansen, I.E., Sondervan, D., Seelaar, H., Blake, D., Young, K., Halliwell, N., Callister, J.B., Toulson, G., Richardson, A., Gerhard, A., Snowden, J., Mann, D., Neary, D., Nalls, M.A., Peuralinna, T., Jansson, L., Isoviita, V.M., Kaivorinne, A.L., Hölttä-Vuori, M., Ikonen, E., Sulkava, R., Benatar, M., Wuu, J., Chiò, A., Restagno, G., Borghero, G., Sabatelli, M., Heckerman, D., Rogaeva, E., Zinman, L., Rothstein, J.D., Sendtner, M., Drepper, C., Eichler, E.E., Alkan, C., Abdullaev, Z., Pack, S.D., Dutra, A., Pak, E., Hardy, J., Singleton, A., Williams, N.M., Heutink, P., Pickering-Brown, S., Morris, H.R., Tienari, P.J., Traynor, B.J., 2011. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 72, 257–268.

Ritson, G.P., Custer, S.K., Freibaum, B.D., Guinto, J.B., Geffel, D., Moore, J., Tang, W., Winton, M.J., Neumann, M., Trojanowski, J.Q., Lee, V.M.-Y., Forman, M.S., Taylor, J.P., 2010. TDP-43 Mediates Degeneration in a Novel Drosophila Model of Disease Caused by Mutations in VCP/p97. J. Neurosci. 30, 7729–7739.

Robberecht, W., Philips, T., 2013. The changing scene of amyotrophic lateral sclerosis. Nat. Rev. Neurosci. 14, 248–264.

Robinson, A., Davidson, Y., Snowden, J.S., Mann, D.M.A., 2014. C9ORF72 in Dementia with Lewy bodies. J. Neurol. Neurosurg. Psychiatry 85, 10–12.

Rogers, D.W., Böttcher, M.A., Traulsen, A., Greig, D., 2017. Ribosome reinitiation can explain length-dependent translation of messenger RNA. PLoS Comput. Biol. 13, e1005592.

Rohrer, J.D., Isaacs, A.M., Mizlienska, S., Mead, S., Lashley, T., Wray, S., Sidle, K., Fratta, P., Orrell, R.W., Hardy, J., Holton, J., Revesz, T., Rossor, M.N., Warren, J.D., 2015. C9orf72 expansions in frontotemporal dementia and amyotrophic lateral sclerosis. Lancet Neurol.

Romano, M., Buratti, E., Romano, G., Klima, R., Del Bel Belluz, L., Stuani, C., Baralle, F., Feiguin, F., 2014. Evolutionarily conserved heterogeneous nuclear ribonucleoprotein (hnRNP) A/B proteins functionally interact with human and drosophila tar DNA-binding protein 43 (TDP-43). J. Biol. Chem. 289, 7121–7130.

Roussos, P., Katsel, P., Davis, K.L., Giakoumaki, S.G., Siever, L.J., Bitsios, P., Haroutunian, V., 2013. Convergent Findings for Abnormalities of the NF-κB Signaling Pathway in Schizophrenia. Neuropsychopharmacology 38, 533–539.

Rubinsztein, D.C., Bento, C.F., Deretic, V., 2015. Therapeutic targeting of autophagy in neurodegenerative and infectious diseases. J. Exp. Med. 212, 979–990.

Rudich, P., Snoznik, C., Watkins, S.C., Monaghan, J., Bhan Pandey, U., Lamitina, T., 2017. Nuclear localized C9orf72 associated arginine containing dipeptides exhibit age-dependent toxicity in C. elegans. Hum. Mol. Genet.

Russ, J., Liu, E.Y., Wu, K., Neal, D., Suh, E.R., Irwin, D.J., McMillan, C.T., Harms, M.B., Cairns, N.J., Wood, E.M., Xie, S.X., Elman, L., McCluskey, L., Grossman, M., Van Deerlin, V.M., Lee, E.B., 2015. Hypermethylation of repeat expanded C9orf72 is a clinical and molecular disease modifier. Acta Neuropathol. 129, 39–52.

Russo, A., Scardigli, R., Regina, F. La, Murray, M.E., Romano, N., Dickson, D.W., Wolozin, B., Cattaneo, A., Ceci, M., 2017. Increased cytoplasmic TDP-43 reduces global protein synthesis by interacting with Rack1 on polyribosomes. Hum. Mol. Genet. 26, 1407–1418.

Rusten, T.E., Stenmark, H., 2010. p62, an autophagy hero or culprit? Nat. Cell Biol. 12, 207–209.

Rutherford, N.J., Heckman, M.G., DeJesus-Hernandez, M., Baker, M.C., Soto-Ortolaza, A.I., Rayaprolu, S., Stewart, H., Finger, E., Volkening, K., Seeley, W.W., Hatanpaa, K.J., Lomen-Hoerth, C., Kertesz, A., Bigio, E.H., Lippa, C., Knopman, D.S., Kretzschmar, H.A., Neumann, M., Caselli, R.J., White, C.L.,

Mackenzie, I.R., Petersen, R.C., Strong, M.J., Miller, B.L., Boeve, B.F., Uitti, R.J., Boylan, K.B., Wszolek, Z.K., Graff-Radford, N.R., Dickson, D.W., Ross, O.A., Rademakers, R., 2012. Length of normal alleles of C9ORF72 GGGGCC repeat do not influence disease phenotype. Neurobiol. Aging 33, 2950.e5-2950.e7.

Saberi S, Stauffer JE, Jiang J, Garcia SD, Taylor AE, Schulte D, et al. Sense-encoded poly-GR dipeptide repeat proteins correlate to neurodegeneration and uniquely co-localize with TDP-43 in dendrites of repeat-expanded C9orf72 amyotrophic lateral sclerosis. Acta Neuropathol 2018; 135: 459-474

Saez, I., Vilchez, D., 2014. The Mechanistic Links Between Proteasome Activity, Aging and Agerelated Diseases. Curr. Genomics 15, 38–51.

Saitoh, T., Fujita, N., Jang, M.H., Uematsu, S., Yang, B.-G., Satoh, T., Omori, H., Noda, T., Yamamoto, N., Komatsu, M., Tanaka, K., Kawai, T., Tsujimura, T., Takeuchi, O., Yoshimori, T., Akira, S., 2008. Loss of the autophagy protein Atg16L1 enhances endotoxin-induced IL-1β production. Nature 456, 264–268.

Sang, J.L., 1962. The nucleolus. Progr. Biophys. Chem. 12, 25–66.

Santoro, M.R., Bray, S.M., Warren, S.T., 2012. Molecular Mechanisms of Fragile X Syndrome: A Twenty-Year Perspective. Annu. Rev. Pathol. Mech. Dis. 7, 219–245.

Sareen, D., O'Rourke, J.G., Meera, P., Muhammad, A.K.M.G., Grant, S., Simpkinson, M., Bell, S., Carmona, S., Ornelas, L., Sahabian, A., Gendron, T., Petrucelli, L., Baughn, M., Ravits, J., Harms, M.B., Rigo, F., Bennett, C.F., Otis, T.S., Svendsen, C.N., Baloh, R.H., 2013. Targeting RNA Foci in iPSC-Derived Motor Neurons from ALS Patients with a C9ORF72 Repeat Expansion. Sci. Transl. Med. 5, 208ra149-208ra149.

Sasaki, S., Goldberg, A., Klionsky, D., Mizushima, N., Ohsumi, Y., Yoshimori, T., Levine, B., Kroemer, G., Baehrecke, E., Shintani, T., Klionsky, D., Borsello, T., Croquelois, K., Hornung, J., Levine, B., Klionsky, D., Cuervo, A., Bergamini, E., Brunk, U., Kondo, Y., Kanazawa, T., Sawaya, R., Chu, C., Mizushima, N., Levine, B., Cuervo, A., Hara, T., Nakamura, K., Matsui, M., Komatsu, M., Waguri, S., Chiba, T., Nixon, R., Rubinsztein, D., Nakano, I., Shibata, T., Uesaka, Y., Morimoto, N., Nagai, M., Ohta, Y., Li, L., Zhang, X., Le, W., Rusten, T., Simonsen, A., Kabeya, Y., Mizushima, N., Ueno, T., Mizushima, N., Yamamoto, A., Matsui, M., Ichimura, Y., Kumamomidou, T., Sou, Y., Sasaki, S., Iwata, M., Tsukada, M., Ohsumi, Y., Kuma, A., Hatano, M., Matsui, M., Iwata, J., Ezaki, J., Komatsu, M., Elmore, S., Qian, T., Grissom, S., Fortun, J., Dunn, W., Joy, S., Ravikumar, B., Vacher, C., Berger, Z., Ravikumar, B., Duden, R., Rubinsztein, D., Iwata, A., Christianson, J., Bucci, M., Komatsu, M., Waguri, S., Koike, M., Yue, Z., Horton, A., Bravin, M., Koike, M., Tadakoshi, M., Gotoh, K., Kohli, L., Roth, K., Qin, Z., Wang, Y., Kegel, K., Anglade, P., Vyas, S., Javoy-Agid, F., Nixon, R., Wegiel, J., Kumar, A., Liberski, P., Sikorska, B., Bratosiewicz-Wasik, J., Lee, J., Gao, F., Wang, Q., Ding, Y., Kohtz, D., Bjørkøy, G., Lamark, T., Brech, A., Pankiv, S., Clausen, T., Lamark, T., De, D.C.,

Wattiaux, R., Webb, J., Ravikumar, B., Atkins, J., Fornai, F., Longone, P., Cafaro, L., Aguib, Y., Heiseke, A., Gilch, S., Mizuno, Y., Amari, M., Takatama, M., Sasaki, S., Hirano, A., Sax, D., Zimmermann, H., Sotelo, C., Palay, S., Sasaki, S., Hirano, A., Donnenfeld, H., Sarkar, S., Floto, R., Berger, Z., Williams, A., Sarkar, S., Cuddon, P., Levine, B., Yuan, J., Mizushima, N., Hollenbeck, P., Yue, Z., Pasquali, L., Longone, P., Isidoro, C., 2011. Autophagy in spinal cord motor neurons in sporadic amyotrophic lateral sclerosis. J. Neuropathol. Exp. Neurol. 70, 349–59.

Sato, N., Amino, T., Kobayashi, K., Asakawa, S., Ishiguro, T., Tsunemi, T., Takahashi, M., Matsuura, T., Flanigan, K.M., Iwasaki, S., Ishino, F., Saito, Y., Murayama, S., Yoshida, M., Hashizume, Y., Takahashi, Y., Tsuji, S., Shimizu, N., Toda, T., Ishikawa, K., Mizusawa, H., 2009. Spinocerebellar Ataxia Type 31 Is Associated with “Inserted” Penta-Nucleotide Repeats Containing (TGGAA)n. Am. J. Hum. Genet. 85, 544–557.

Savas, J.N., Toyama, B.H., Xu, T., Yates, J.R., Hetzer, M.W., 2012. Extremely Long-Lived Nuclear Pore Proteins in the Rat Brain. Science (80-.). 335, 942–942.

Scaber, J., Talbot, K., 2016. What is the role of TDP-43 in C9orf72-related amyotrophic lateral sclerosis and frontemporal dementia? Brain. 139, 3057-3059.

Schleich, S., Strassburger, K., Janiesch, P.C., Koledachkina, T., Miller, K.K., Haneke, K., Cheng, Y.-S., Küchler, K., Stoecklin, G., Duncan, K.E., Teleman, A.A., 2014. DENR–MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth. Nature 512, 208–212.

Schludi, M.H., Becker, L., Garrett, L., Gendron, T.F., Zhou, Q., Schreiber, F., Popper, B., Dimou, L., Strom, T.M., Winkelmann, J., von Thaden, A., Rentzsch, K., May, S., Michaelsen, M., Schwenk, B.M., Tan, J., Schoser, B., Dieterich, M., Petrucelli, L., Hölter, S.M., Wurst, W., Fuchs, H., Gailus-Durner, V., de Angelis, M.H., Klopstock, T., Arzberger, T., Edbauer, D., 2017. Spinal poly-GA inclusions in a C9orf72 mouse model trigger motor deficits and inflammation without neuron loss. Acta Neuropathol. 134, 241–254.

Schludi, M.H., May, S., Grässer, F.A., Rentzsch, K., Kremmer, E., Küpper, C., Klopstock, T., Ceballos-Baumann, A., Danek, A., Diehl-Schmid, J., Fassbender, K., Förstl, H., Kornhuber, J., Otto, M., Dieterich, M., Feuerecker, R., Giese, A., Klünemann, H., Kurz, A., Levin, J., Lorenzl, S., Meyer, T., Nübling, G., Roeber, S., 2015. Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. Acta Neuropathol. 130, 537–555.

Schmid, B., Hruscha, A., Hogl, S., Banzhaf-Strathmann, J., Strecker, K., van der Zee, J., Teucke, M., Eimer, S., Hegermann, J., Kittelmann, M., Kremmer, E., Cruts, M., Solchenberger, B., Hasenkamp, L., van Bebber, F., Van Broeckhoven, C., Edbauer, D., Lichtenthaler, S.F., Haass, C., 2013. Loss of ALS-associated TDP-43 in zebrafish causes muscle degeneration , vascular

dysfunction , and reduced motor neuron axon outgrowth. Proc. Natl. Acad. Sci. U. S. A. 110, 1–6.

Schmidt, H.B., Gorlich, D., 2016. Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. Trends Biochem. Sci.

Schmidt, H.B., Rohatgi, R., 2016. In vivo Formation of Vacuolated Multi-phase Compartments Lacking Membranes. Cell Rep. 16, 1228–1236.

Scotter, E.L., Chen, H.J., Shaw, C.E., 2015. TDP-43 Proteinopathy and ALS: Insights into Disease Mechanisms and Therapeutic Targets. Neurotherapeutics.

Seelaar, H., Jurgen Schelhaas, H., Azmani, A., Küsters, B., Rosso, S., Majoor-Krakauer, D., De Rijik, M.C., Rizzu, P., Ten Brummelhuis, M., Van Doorn, P.A., Kamphorst, W., Willemsen, R., Van Swieten, J.C., 2007. TDP-43 pathology in familial frontotemporal dementia and motor neuron disease without Progranulin mutations. Brain 130, 1375–1385.

Seelaar, H., Rohrer, J.D., Pijnenburg, Y.A.L., Fox, N.C., van Swieten, J.C., 2011. Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a review. J. Neurol. Neurosurg. Psychiatry 82, 476–486.

Sellier, C., Buijsen, R.A.M., He, F., Natla, S., Jung, L., Tropel, P., Gaucherot, A., Jacobs, H., Meziane, H., Vincent, A., Champy, M.F., Sorg, T., Pavlovic, G., Wattenhofer-Donze, M., Birling, M.C., Oulad-Abdelghani, M., Eberling, P., Ruffenach, F., Joint, M., Anheim, M., Martinez-Cerdeno, V., Tassone, F., Willemsen, R., Hukema, R.K., Viville, S., Martinat, C., Todd, P.K., Charlet-Berguerand, N., 2017. Translation of Expanded CGG Repeats into FMRpolyG Is Pathogenic and May Contribute to Fragile X Tremor Ataxia Syndrome. Neuron 93, 331–347.

Sellier, C., Campanari, M., Julie Corbier, C., Gaucherot, A., Kolb-Cheynel, I., Oulad-Abdelghani, M., Ruffenach, F., Page, A., Ciura, S., Kabashi, E., Charlet-Berguerand, N., 2016. Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. EMBO J. 35, 1276–1297.

Shatunov, A., Mok, K., Newhouse, S., Weale, M.E., Smith, B., Vance, C., Johnson, L., Veldink, J.H., van Es, M.A., van den Berg, L.H., Robberecht, W., Van Damme, P., Hardiman, O., Farmer, A.E., Lewis, C.M., Butler, A.W., Abel, O., Andersen, P.M., Fogh, I., Silani, V., Chiò, A., Traynor, B.J., Melki, J., Meininger, V., Landers, J.E., McGuffin, P., Glass, J.D., Pall, H., Leigh, P.N., Hardy, J., Brown, R.H., Powell, J.F., Orrell, R.W., Morrison, K.E., Shaw, P.J., Shaw, C.E., Al-Chalabi, A., 2010. Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: A genome-wide association study. Lancet Neurol. 9, 986–994.

Shi, Y., Lin, S., Staats, K.A., Li, Y., Chang, W.H., Hung, S.T., Hendricks, E., Linares, G.R., Wang, Y., Son, E.Y., Wen, X., Kisler, K., Wilkinson, B., Menendez, L., Sugawara, T., Woolwine, P., Huang, M., Cowan, M.J., Ge, B., Koutsodendris,

N., Sandor, K.P., Komberg, J., Vangoor, V.R., Senthilkumar, K., Hennes, V., Seah, C., Nelson, A.R., Cheng, T.Y., Lee, S.J.J., August, P.R., Chen, J.A., Wisniewski, N., Hanson-Smith, V., Belgard, T.G., Zhang, A., Coba, M., Grunseich, C., Ward, M.E., Van Den Berg, L.H., Pasterkamp, R.J., Trotti, D., Zlokovic, B. V, Ichida, J.K., 2018. Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. Nat. Med. 24, 313–325..

Shi, K.Y., Mori, E., Nizami, Z.F., Lin, Y., Kato, M., Xiang, S., Wu, L.C., Ding, M., Yu, Y., Gall, J.G., McKnight, S.L., 2017. Toxic PR n poly-dipeptides encoded by the C9orf72 repeat expansion block nuclear import and export. Proc. Natl. Acad. Sci. 114, E1111–E1117.

Shi, W.Y., Skeath, J.B., 2004. The Drosophila RCC1 homolog, Bj1, regulates nucleocytoplasmic transport and neural differentiation during Drosophila development. Dev. Biol. 270, 106–121.

Simón-Sánchez, J., Dopper, E.G.P., Cohn-Hokke, P.E., Hukema, R.K., Nicolaou, N., Seelaar, H., De Graaf, J.R.A., De Koning, I., Van Schoor, N.M., Deeg, D.J.H., Smits, M., Raaphorst, J., Van Den Berg, L.H., Schelhaas, H.J., De Die-Smulders, C.E.M., Majoor-Krakauer, D., Rozemuller, A.J.M., Willemsen, R., Pijnenburg, Y.A.L., Heutink, P., Van Swieten, J.C., 2012. The clinical and pathological phenotype of C9ORF72 hexanucleotide repeat expansions. Brain 135, 723–735.

Sivadasan, R., Hornburg, D., Drepper, C., Frank, N., Jablonka, S., Hansel, A., Lojewski, X., Sterneckert, J., Hermann, A., Shaw, P.J., Ince, P.G., Mann, M., Meissner, F., Sendtner, M., 2016. C9ORF72 interaction with cofilin modulates actin dynamics in motor neurons. Nat. Neurosci. 19, 1610–1618.

Smith, B.N., Ticozzi, N., Fallini, C., Gkazi, A.S., Topp, S., Kenna, K.P., Scotter, E.L., Kost, J., Keagle, P., Miller, J.W., Calini, D., Vance, C., Danielson, E.W., Troakes, C., Tiloca, C., Al-Sarraj, S., Lewis, E.A., King, A., Colombrita, C., Pensato, V., Castellotti, B., de Belleroche, J., Baas, F., ten Asbroek, A.L., Sapp, P.C., McKenna-Yasek, D., McLaughlin, R.L., Polak, M., Asress, S., Esteban-Pérez, J., Muñoz-Blanco, J.L., Simpson, M., van Rheenen, W., Diekstra, F.P., Lauria, G., Duga, S., Corti, S., Cereda, C., Corrado, L., Sorarù, G., Morrison, K.E., Williams, K.L., Nicholson, G.A., Blair, I.P., Dion, P.A., Leblond, C.S., Rouleau, G.A., Hardiman, O., Veldink, J.H., Van Den Berg, L.H., Al-Chalabi, A., Pall, H., Shaw, P.J., Turner, M.R., Talbot, K., Taroni, F., García-Redondo, A., Wu, Z., Glass, J.D., Gellera, C., Ratti, A., Brown, R.H., Silani, V., Shaw, C.E., Landers, J.E., 2014. Exome-wide rare variant analysis identifies TUBA4A mutations associated with familial ALS. Neuron 84, 324–331.

Smith, B.N., Topp, S.D., Fallini, C., Shibata, H., Chen, H.-J., Troakes, C., King, A., Ticozzi, N., Kenna, K.P., Soragia-Gkazi, A., Miller, J.W., Sato, A., Dias, D.M., Jeon, M., Vance, C., Wong, C.H., de Majo, M., Kattuah, W., Mitchell, J.C., Scotter, E.L., Parkin, N.W., Sapp, P.C., Nolan, M., Nestor, P.J., Simpson, M., Weale, M., Lek, M., Baas, F., Vianney de Jong, J.M., ten Asbroek, A.L.M.A., Redondo, A.G., Esteban-Pérez, J., Tiloca, C., Verde, F., Duga, S., Leigh, N., Pall, H., Morrison, K.E., Al-Chalabi, A., Shaw, P.J., Kirby, J., Turner, M.R.,

Talbot, K., Hardiman, O., Glass, J.D., De Belleroche, J., Maki, M., Moss, S.E., Miller, C., Gellera, C., Ratti, A., Al-Sarraj, S., Brown, R.H., Silani, V., Landers, J.E., Shaw, C.E., 2017. Mutations in the vesicular trafficking protein annexin A11 are associated with amyotrophic lateral sclerosis. Sci. Transl. Med. 9, eaad9157.

Smith, B.N., Newhouse, S., Shatunov, A., Vance, C., Topp, S., Johnson, L., Miller, J., Lee, Y., Troakes, C., Scott, K.M., Jones, A., Gray, I., Wright, J., Hortobágyi, T., Al-Sarraj, S., Rogelj, B., Powell, J., Lupton, M., Lovestone, S., Sapp, P.C., Weber, M., Nestor, P.J., Schelhaas, H.J., Asbroek, A.A. ten, Silani, V., Gellera, C., Taroni, F., Ticozzi, N., Van den Berg, L., Veldink, J., Van Damme, P., Robberecht, W., Shaw, P.J., Kirby, J., Pall, H., Morrison, K.E., Morris, A., de Belleroche, J., Vianney de Jong, J.M.B., Baas, F., Andersen, P.M., Landers, J., Brown, R.H., Weale, M.E., Al-Chalabi, A., Shaw, C.E., 2013. The C9ORF72 expansion mutation is a common cause of ALS+/−FTD in Europe and has a single founder. Eur. J. Hum. Genet. 21, 102–108.

Snowden, J.S., Rollinson, S., Thompson, J.C., Harris, J.M., Stopford, C.L., Richardson, A.M.T., Jones, M., Gerhard, A., Davidson, Y.S., Robinson, A., Gibbons, L., Hu, Q., DuPlessis, D., Neary, D., Mann, D.M.A., Pickering-Brown, S.M., 2012. Distinct clinical and pathological characteristics of frontotemporal dementia associated with C9ORF72 mutations. Brain 135, 693–708.

Song, J., Perreault, J.-P., Topisirovic, I., Richard, S., 2016. RNA G-quadruplexes and their potential regulatory roles in translation. Translation 4, e1244031.

Spector, D.L., Lamond, A.I., 2011. Nuclear speckles. Cold Spring Harb. Perspect. Biol. 3, 1–12.

St Johnston, D., 2002. The art and design of genetic screens: Drosophila Melanogaster. Nat. Rev. Genet. 3, 176–188.

Štalekar, M., Yin, X., Rebolj, K., Darovic, S., Troakes, C., Mayr, M., Shaw, C.E., Rogelj, B., 2015. Proteomic analyses reveal that loss of TDP-43 affects RNA processing and intracellular transport. Neuroscience 293, 157–170.

Stan, D.A., Subroto, G., Gao, X.M., Roberts, C.R., Lewis-Amezcua, K., Hatanpaa, J.K., Tamminga, A.C., 2006. Human Postmortem Tissue: What Quality Markers Matter?. J. Brain Res. 1123, 1–11.

Stenmark, H., 2009. Rab GTPases as coordinators of vesicle traffic. Nat. Rev. Mol. Cell Biol. 10, 513–525.

Stepto, A., Gallo, J.M., Shaw, C.E., Hirth, F., 2014. Modelling C9ORF72 hexanucleotide repeat expansion in amyotrophic lateral sclerosis and frontotemporal dementia. Acta Neuropathol.

Stewart, M., 2007. Molecular mechanism of the nuclear protein import cycle. Nat. Rev. Mol. Cell Biol. 8, 195–208.

Strong, M.J., Yang, W., 2011. The frontotemporal syndromes of ALS. Clinicopathological correlates. In: Journal of Molecular Neuroscience. Humana Press Inc, pp. 648–655.

Su, Z., Zhang, Y., Gendron, T.F., Bauer, P.O., Chew, J., Yang, W.Y., Fostvedt, E., Jansen-West, K., Belzil, V. V., Desaro, P., Johnston, A., Overstreet, K., Oh, S.Y., Todd, P.K., Berry, J.D., Cudkowicz, M.E., Boeve, B.F., Dickson, D., Floeter, M.K., Traynor, B.J., Morelli, C., Ratti, A., Silani, V., Rademakers, R., Brown, R.H., Rothstein, J.D., Boylan, K.B., Petrucelli, L., Disney, M.D., 2014. Discovery of a Biomarker and Lead Small Molecules to Target r(GGGGCC)-Associated Defects in c9FTD/ALS. Neuron 83, 1043–1050.

Subramanian, M., Rage, F., Tabet, R., Flatter, E., Mandel, J.-L., Moine, H., 2011. G–quadruplex RNA structure as a signal for neurite mRNA targeting. EMBO Rep. 12, 697–704.

Sudria-Lopez, E., Koppers, M., de Wit, M., van der Meer, C., Westeneng, H.J., Zundel, C.A.C., Youssef, S.A., Harkema, L., de Bruin, A., Veldink, J.H., van den Berg, L.H., Pasterkamp, R.J., 2016. Full ablation of C9orf72 in mice causes immune system-related pathology and neoplastic events but no motor neuron defects. Acta Neuropathol.

Suh, E.R., Lee, E.B., Neal, D., Wood, E.M., Toledo, J.B., Rennert, L., Irwin, D.J., McMillan, C.T., Krock, B., Elman, L.B., McCluskey, L.F., Grossman, M., Xie, S.X., Trojanowski, J.Q., Van Deerlin, V.M., 2015. Semi-automated quantification of C9orf72 expansion size reveals inverse correlation between hexanucleotide repeat number and disease duration in frontotemporal degeneration. Acta Neuropathol. 130, 363–372.

Suhonen, N.M., Kaivorinne, A.L., Moilanen, V., Bode, M., Takalo, R., Hanninen, T., Remes, A.M., 2015. Slowly progressive frontotemporal lobar degeneration caused by the C9ORF72 repeat expansion: a 20-year follow-up study. Neurocase 21, 85–89.

Sullivan, P.M., Zhou, X., Robins, A.M., Paushter, D.H., Kim, D., Smolka, M.B., Hu, F., 2016. The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. Acta Neuropathol. Commun. 4, 51.

Sun, Y., Chakrabartty, A., 2017. Phase to Phase with TDP-43. Biochemistry. 56, 809–823

Suzuki, N., Maroof, A.M., Merkle, F.T., Koszka, K., Intoh, A., Armstrong, I., Moccia, R., Davis-Dusenbery, B.N., Eggan, K., 2013. The mouse C9ORF72 ortholog is enriched in neurons known to degenerate in ALS and FTD. Nat. Neurosci. 16, 1725–1727.

Swarup, V., Phaneuf, D., Dupré, N., Petri, S., Strong, M., Kriz, J., Julien, J.-P., 2011. Deregulation of TDP-43 in amyotrophic lateral sclerosis triggers nuclear factor κB–mediated pathogenic pathways. J. Exp. Med. 208, 2429–2447.

Swindell, W.R., Masternak, M.M., Kopchick, J.J., Conover, C.A., Bartke, A., Miller, R.A., 2009. Endocrine regulation of heat shock protein mRNA levels in long-lived dwarf mice. Mech. Ageing Dev. 130, 393–400.

Swinnen, B., Bento-Abreu, A., Gendron, T.F., Boeynaems, S., Bogaert, E., Nuyts, R., Timmers, M., Scheveneels, W., Hersmus, N., Wang, J., Mizielinska, S., Isaacs, A.M., Petrucelli, L., Lemmens, R., Van Damme, P., Van Den Bosch, L., Robberecht, W., 2018. A zebrafish model for C9orf72 ALS reveals RNA toxicity as a pathogenic mechanism. Acta Neuropathol. 135, 427–443.

Swinnen, B., Robberecht, W., 2014. The phenotypic variability of amyotrophic lateral sclerosis. Nat. Rev. Neurol. 10, 661–670.

Tabet R, Schaeffer L, Freyermuth F, Jambeau M, Workman M, Lee CZ, et al. CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD C9ORF72 transcripts. Nat Commun 2018; 9: 152

Takeuchi, R., Toyoshima, Y., Tada, M., Shiga, A., Tanaka, H., Shimohata, M., Kimura, K., Morita, T., Kakita, A., Nishizawa, M., Takahashi, H., 2013. Transportin 1 accumulates in FUS inclusions in adult-onset ALS without FUS mutation. Neuropathol. Appl. Neurobiol.

Talbot, K., 2009. Motor neuron disease: the bare essentials. Pract. Neurol. 9, 303–309.

Tao, Z., Wang, H., Xia, Q., Li, K., Li, K., Jiang, X., Xu, G., Wang, G., Ying, Z., 2015. Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity. Hum. Mol. Genet. 24, 2426–2441.

Tartaglia, M.C., Rowe, A., Findlater, K., Orange, J.B., Grace, G., Strong, M.J., 2007. Differentiation between primary lateral sclerosis and amyotrophic lateral sclerosis: examination of symptoms and signs at disease onset and during follow-up. Arch. Neurol. 64, 232–236.

Taylor, J.P., Brown, R.H., Cleveland, D.W., 2016. Decoding ALS: from genes to mechanism. Nature 539, 197–206.

Tekotte, H., Berdnik, D., Török, T., Buszczak, M., Jones, L.M., Cooley, L., Knoblich, J.A., Davis, I., 2002. Dcas is required for importin-alpha3 nuclear export and mechano-sensory organ cell fate specification in Drosophila. Dev. Biol. 244, 396–406.

Therrien, M., Rouleau, G.A., Dion, P.A., Parker, J.A., 2013. Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in C. elegans. PLoS One 8, e83450.

Todd, P.K., Oh, S.Y., Krans, A., He, F., Sellier, C., Frazer, M., Renoux, A.J., Chen, K. chun, Scaglione, K.M., Basrur, V., Elenitoba-Johnson, K., Vonsattel, J.P., Louis, E.D., Sutton, M.A., Taylor, J.P., Mills, R.E., Charlet-Berguerand, N.,

Paulson, H.L., 2013. CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. Neuron 78, 440–455.

Toyama, B.H., Savas, J.N., Park, S.K., Harris, M.S., Ingolia, N.T., Yates, J.R., Hetzer, M.W., 2013. Identification of long-lived proteins reveals exceptional stability of essential cellular structures. Cell 154, 971–982.

Tran, H., Almeida, S., Moore, J., Gendron, T.F., Chalasani, U.D., Lu, Y., Du, X., Nickerson, J.A., Petrucelli, L., Weng, Z., Gao, F.B., 2015. Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a Drosophila Model of C9ORF72 FTD/ALS. Neuron 87, 1207–1214.

Troakes, C., Maekawa, S., Wijesekera, L., Rogelj, B., Siklós, L., Bell, C., Smith, B., Newhouse, S., Vance, C., Johnson, L., Hortobágyi, T., Shatunov, A., Al-Chalabi, A., Leigh, N., Shaw, C.E., King, A., Al-Sarraj, S., 2012. An MND/ALS phenotype associated with C9orf72 repeat expansion: Abundant p62-positive, TDP-43-negative inclusions in cerebral cortex, hippocampus and cerebellum but without associated cognitive decline. Neuropathology 32, 505–514.

Turner, M.R., Al-Chalabi, A., Chio, A., Hardiman, O., Kiernan, M.C., Rohrer, J.D., Rowe, J., Seeley, W., Talbot, K., 2017. Genetic screening in sporadic ALS and FTD. J. Neurol. Neurosurg. Psychiatry jnnp-2017-315995.

Twiss, J.L., Kalinski, A.L., Sachdeva, R., Houle, J.D., 2016. Intra-axonal protein synthesis – A new target for neural repair? Neural Regen. Res.

Udan-Johns, M., Bengoechea, R., Bell, S., Shao, J., Diamond, M.I., True, H.L., Weihl, C.C., Baloh, R.H., 2014. Prion-like nuclear aggregation of TDP-43 during heat shock is regulated by HSP40/70 chaperones. Hum. Mol. Genet. 23, 157–170.

Ugolino, J., Ji, Y.J., Conchina, K., Chu, J., Nirujogi, R.S., Pandey, A., Brady, N.R., Hamacher-Brady, A., Wang, J., 2016. Loss of C9orf72 Enhances Autophagic Activity via Deregulated mTOR and TFEB Signaling. PLoS Genet. 12, e1006443.

van Blitterswijk, M., DeJesus-Hernandez, M., Niemantsverdriet, E., Murray, M.E., Heckman, M.G., Diehl, N.N., Brown, P.H., Baker, M.C., Finch, N.C.A., Bauer, P.O., Serrano, G., Beach, T.G., Josephs, K.A., Knopman, D.S., Petersen, R.C., Boeve, B.F., Graff-Radford, N.R., Boylan, K.B., Petrucelli, L., Dickson, D.W., Rademakers, R., 2013. Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): A cross-sectional cohort study. Lancet Neurol. 12, 978–988.

van Blitterswijk, M., Gendron, T.F., Baker, M.C., DeJesus-Hernandez, M., Finch, N.C.A., Brown, P.H., Daughrity, L.M., Murray, M.E., Heckman, M.G., Jiang, J., Lagier-Tourenne, C., Edbauer, D., Cleveland, D.W., Josephs, K.A., Parisi, J.E., Knopman, D.S., Petersen, R.C., Petrucelli, L., Boeve, B.F., Graff-Radford, N.R., Boylan, K.B., Dickson, D.W., Rademakers, R., 2015. Novel clinical associations

with specific C9ORF72 transcripts in patients with repeat expansions in C9ORF72. Acta Neuropathol. 130, 863–876.

van Blitterswijk, M., Mullen, B., Heckman, M.G., Baker, M.C., DeJesus-Hernandez, M., Brown, P.H., Murray, M.E., Hsiung, G.Y.R., Stewart, H., Karydas, A.M., Finger, E., Kertesz, A., Bigio, E.H., Weintraub, S., Mesulam, M., Hatanpaa, K.J., White, C.L., Neumann, M., Strong, M.J., Beach, T.G., Wszolek, Z.K., Lippa, C., Caselli, R., Petrucelli, L., Josephs, K.A., Parisi, J.E., Knopman, D.S., Petersen, R.C., Mackenzie, I.R., Seeley, W.W., Grinberg, L.T., Miller, B.L., Boylan, K.B., Graff-Radford, N.R., Boeve, B.F., Dickson, D.W., Rademakers, R., 2014. Ataxin-2 as potential disease modifier in C9ORF72 expansion carriers. Neurobiol. Aging 35, 2421.e13-7.

Van Deerlin, V.M., Sleiman, P.M.A., Martinez-Lage, M., Chen-Plotkin, A., Wang, L.-S., Graff-Radford, N.R., Dickson, D.W., Rademakers, R., Boeve, B.F., Grossman, M., Arnold, S.E., Mann, D.M.A., Pickering-Brown, S.M., Seelaar, H., Heutink, P., van Swieten, J.C., Murrell, J.R., Ghetti, B., Spina, S., Grafman, J., Hodges, J., Spillantini, M.G., Gilman, S., Lieberman, A.P., Kaye, J.A., Woltjer, R.L., Bigio, E.H., Mesulam, M., Al-Sarraj, S., Troakes, C., Rosenberg, R.N., White, C.L., Ferrer, I., Lladó, A., Neumann, M., Kretzschmar, H.A., Hulette, C.M., Welsh-Bohmer, K.A., Miller, B.L., Alzualde, A., de Munain, A.L., McKee, A.C., Gearing, M., Levey, A.I., Lah, J.J., Hardy, J., Rohrer, J.D., Lashley, T., Mackenzie, I.R.A., Feldman, H.H., Hamilton, R.L., Dekosky, S.T., van der Zee, J., Kumar-Singh, S., Van Broeckhoven, C., Mayeux, R., Vonsattel, J.P.G., Troncoso, J.C., Kril, J.J., Kwok, J.B.J., Halliday, G.M., Bird, T.D., Ince, P.G., Shaw, P.J., Cairns, N.J., Morris, J.C., McLean, C.A., DeCarli, C., Ellis, W.G., Freeman, S.H., Frosch, M.P., Growdon, J.H., Perl, D.P., Sano, M., Bennett, D.A., Schneider, J.A., Beach, T.G., Reiman, E.M., Woodruff, B.K., Cummings, J., Vinters, H. V, Miller, C.A., Chui, H.C., Alafuzoff, I., Hartikainen, P., Seilhean, D., Galasko, D., Masliah, E., Cotman, C.W., Tuñón, M.T., Martínez, M.C.C., Munoz, D.G., Carroll, S.L., Marson, D., Riederer, P.F., Bogdanovic, N., Schellenberg, G.D., Hakonarson, H., Trojanowski, J.Q., Lee, V.M.-Y., 2010. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. Nat. Genet. 42, 234–239.

van der Zee, J., Gijselinck, I., Dillen, L., Van Langenhove, T., Theuns, J., Engelborghs, S., Philtjens, S., Vandenbulcke, M., Sleegers, K., Sieben, A., Bäumer, V., Maes, G., Corsmit, E., Borroni, B., Padovani, A., Archetti, S., Perneczky, R., Diehl-Schmid, J., de Mendonça, A., Miltenberger-Miltenyi, G., Pereira, S., Pimentel, J., Nacmias, B., Bagnoli, S., Sorbi, S., Graff, C., Chiang, H.H., Westerlund, M., Sanchez-Valle, R., Llado, A., Gelpi, E., Santana, I., Almeida, M.R., Santiago, B., Frisoni, G., Zanetti, O., Bonvicini, C., Synofzik, M., Maetzler, W., vom Hagen, J.M., Schöls, L., Heneka, M.T., Jessen, F., Matej, R., Parobkova, E., Kovacs, G.G., Ströbel, T., Sarafov, S., Tournev, I., Jordanova, A., Danek, A., Arzberger, T., Fabrizi, G.M., Testi, S., Salmon, E., Santens, P., Martin, J.J., Cras, P., Vandenberghe, R., De Deyn, P.P., Cruts, M., Van Broeckhoven, C., Ramirez, A., Kurzwelly, D., Sachtleben, C., Mairer, W., Firmo, C., Antonell, A., Molinuevo, J., Forsell, C., Lillius, L., Kinhult Ståhlbom, A., Thonberg, H., Nennesmo, I., Börjesson-Hanson, A., Bessi, V., Piaceri, I., Helena Ribeiro, M., Oliveira, C., Massano, J., Garret, C., Pires, P., Danel, A., Ferrari, S.,

Cavallaro, T., 2013. A Pan-European Study of the C9orf72 Repeat Associated with FTLD: Geographic Prevalence, Genomic Instability, and Intermediate Repeats. Hum. Mutat. 34, 363–373.

Van Mossevelde, S., van der Zee, J., Cruts, M., Van Broeckhoven, C., 2017. Relationship between C9orf72 repeat size and clinical phenotype. Curr. Opin. Genet. Dev. 44, 117-124.

Vance, C., Al-Chalabi, A., Ruddy, D., Smith, B.N., Hu, X., Sreedharan, J., Siddique, T., Schelhaas, H.J., Kusters, B., Troost, D., Baas, F., De Jong, V., Shaw, C.E., 2006. Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2-21.3. Brain 129, 868–876.

Vatsavayai, S.C., Yoon, S.J., Gardner, R.C., Gendron, T.F., Vargas, J.N.S., Trujillo, A., Pribadi, M., Phillips, J.J., Gaus, S.E., Hixson, J.D., Garcia, P.A., Rabinovici, G.D., Coppola, G., Geschwind, D.H., Petrucelli, L., Miller, B.L., Seeley, W.W., 2016. Timing and significance of pathological features in C9orf72 expansion-associated frontotemporal dementia. Brain 139, 3202–3216.

Vernace, V.A., Arnaud, L., Schmidt-Glenewinkel, T., Figueiredo-Pereira, M.E., 2007. Aging perturbs 26S proteasome assembly in Drosophila melanogaster. FASEB J. 21, 2672–2682.

Waite, A.J., Bäumer, D., East, S., Neal, J., Morris, H.R., Ansorge, O., Blake, D.J., 2014. Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. Neurobiol. Aging 35, 1779.e5-1779.e13.

Walker, A.K., Spiller, K.J., Ge, G., Zheng, A., Xu, Y., Zhou, M., Tripathy, K., Kwong, L.K., Trojanowski, J.Q., Lee, V.M.-Y., 2015. Functional recovery in new mouse models of ALS/FTLD after clearance of pathological cytoplasmic TDP-43. Acta Neuropathol. 130, 643–660.

Walker, C., Herranz-Martin, S., Karyka, E., Liao, C., Lewis, K., Elsayed, W., Lukashchuk, V., Chiang, S.-C., Ray, S., Mulcahy, P.J., Jurga, M., Tsagakis, I., Iannitti, T., Chandran, J., Coldicott, I., De Vos, K.J., Hassan, M.K., Higginbottom, A., Shaw, P.J., Hautbergue, G.M., Azzouz, M., El-Khamisy, S.F., 2017. C9orf72 expansion disrupts ATM-mediated chromosomal break repair. Nat. Neurosci. 20, 1225–1235.

Wang, I.F., Wu, L.S., Chang, H.Y., Shen, C.-K.K.J., 2008. TDP-43, the signature protein of FTLD-U, is a neuronal activity-responsive factor. J. Neurochem. 105, 797–806.

Wang, I.-F., Guo, B.-S., Liu, Y.-C., Wu, C.-C., Yang, C.-H., Tsai, K.-J., Shen, C.-K.J., 2012. Autophagy activators rescue and alleviate pathogenesis of a mouse model with proteinopathies of the TAR DNA-binding protein 43. Proc. Natl. Acad. Sci. 109, 15024–15029.

Ward, M.E., Taubes, A., Chen, R., Miller, B.L., Sephton, C.F., Gelfand, J.M., Minami, S., Boscardin, J., Martens, L.H., Seeley, W.W., Yu, G., Herz, J., Filiano, A.J., Arrant, A.E., Roberson, E.D., Kraft, T.W., Farese, R. V., Green, A., Gan, L., 2014. Early retinal neurodegeneration and impaired Ran-mediated nuclear import of TDP-43 in progranulin-deficient FTLD. J. Exp. Med. 211, 1937–1945.

Webster, C.P., Smith, E.F., Bauer, C.S., Moller, A., Hautbergue, G.M., Ferraiuolo, L., Myszczynska, M.A., Higginbottom, A., Walsh, M.J., Whitworth, A.J., Kaspar, B.K., Meyer, K., Shaw, P.J., Grierson, A.J., De Vos, K.J., 2016a. The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. EMBO J. 35, 1656–1676.

Webster, C.P., Smith, E.F., Grierson, A.J., De Vos, K.J., 2016b. C9orf72 plays a central role in Rab GTPase-dependent regulation of autophagy. Small GTPases 1–10.

Weisman-Shomer, P., Cohen, E., Hershco, I., Khateb, S., Wolfovitz-Barchad, O., Hurley, L.H., Fry, M., 2003. The cationic porphyrin TMPyP4 destabilizes the tetraplex form of the fragile X syndrome expanded sequence d(CGG)n. Nucleic Acids Res. 31, 3963–3970.

Wen, X., Tan, W., Westergard, T., Krishnamurthy, K., Markandaiah, S.S., Shi, Y., Lin, S., Shneider, N.A., Monaghan, J., Pandey, U.B., Pasinelli, P., Ichida, J.K., Trotti, D., 2014. Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate invitro and invivo neuronal death. Neuron 84, 1213–1225.

Westergard, T., Jensen, B.K., Wen, X., Cai, J., Kropf, E., Iacovitti, L., Pasinelli, P., Trotti, D., 2016. Cell-to-Cell Transmission of Dipeptide Repeat Proteins Linked to C9orf72-ALS/FTD. Cell Rep. 17, 645–652.

Wheeler, J.R., Matheny, T., Jain, S., Abrisch, R., Parker, R., 2016. Distinct stages in stress granule assembly and disassembly. Elife 5, e18413.

Wilkinson, J.E., Burmeister, L., Brooks, S. V., Chan, C.C., Friedline, S., Harrison, D.E., Hejtmancik, J.F., Nadon, N., Strong, R., Wood, L.K., Woodward, M.A., Miller, R.A., 2012. Rapamycin slows aging in mice. Aging Cell 11, 675–682.

Williams, K.L., Fifita, J.A., Vucic, S., Durnall, J.C., Kiernan, M.C., Blair, I.P., Nicholson, G.A., 2013. Pathophysiological insights into ALS with C9ORF72 expansions. J. Neurol. Neurosurg. Psychiatry 84, 931–935.

Williams, K.L., Topp, S., Yang, S., Smith, B., Fifita, J.A., Warraich, S.T., Zhang, K.Y., Farrawell, N., Vance, C., Hu, X., Chesi, A., Leblond, C.S., Lee, A., Rayner, S.L., Sundaramoorthy, V., Dobson-Stone, C., Molloy, M.P., van Blitterswijk, M., Dickson, D.W., Petersen, R.C., Graff-Radford, N.R., Boeve, B.F., Murray, M.E., Pottier, C., Don, E., Winnick, C., McCann, E.P., Hogan, A., Daoud, H., Levert, A., Dion, P.A., Mitsui, J., Ishiura, H., Takahashi, Y., Goto, J., Kost, J., Gellera, C., Gkazi, A.S., Miller, J., Stockton, J., Brooks, W.S., Boundy, K., Polak, M., Muñoz-Blanco, J.L., Esteban-Pérez, J., Rábano, A., Hardiman, O., Morrison,

K.E., Ticozzi, N., Silani, V., de Belleroche, J., Glass, J.D., Kwok, J.B.J., Guillemin, G.J., Chung, R.S., Tsuji, S., Brown, R.H., García-Redondo, A., Rademakers, R., Landers, J.E., Gitler, A.D., Rouleau, G.A., Cole, N.J., Yerbury, J.J., Atkin, J.D., Shaw, C.E., Nicholson, G.A., Blair, I.P., 2016. CCNF mutations in amyotrophic lateral sclerosis and frontotemporal dementia. Nat. Commun. 7, 11253.

Winton, M.J., Igaz, L.M., Wong, M.M., Kwong, L.K., Trojanowski, J.Q., Lee, V.M.Y., 2008. Disturbance of nuclear and cytoplasmic TAR DNA-binding protein (TDP-43) induces disease-like redistribution, sequestration, and aggregate formation. J. Biol. Chem. 283, 13302–13309.

Wobst, H.J., Wesolowski, S.S., Chadchankar, J., Delsing, L., Jacobsen, S., Mukherjee, J., Deeb, T.Z., Dunlop, J., Brandon, N.J., Moss, S.J., 2017. Cytoplasmic Relocalization of TAR DNA-Binding Protein 43 Is Not Sufficient to Reproduce Cellular Pathologies Associated with ALS In vitro. Front. Mol. Neurosci. 10, 46.

Woerner, A.C., Frottin, F., Hornburg, D., Feng, L.R., Meissner, F., Patra, M., Tatzelt, J., Mann, M., Winklhofer, K.F., Hartl, F.U., Hipp, M.S., 2016. Cytoplasmic protein aggregates interfere with nucleocytoplasmic transport of protein and RNA. Science (80-.). 351, 173–176.

Woollacott, I.O.C., Mead, S., 2014. The C9ORF72 expansion mutation: Gene structure, phenotypic and diagnostic issues. Acta Neuropathol.

Xi, Z., van Blitterswijk, M., Zhang, M., McGoldrick, P., McLean, J.R., Yunusova, Y., Knock, E., Moreno, D., Sato, C., McKeever, P.M., Schneider, R., Keith, J., Petrescu, N., Fraser, P., Tartaglia, M.C., Baker, M.C., Graff-Radford, N.R., Boylan, K.B., Dickson, D.W., Mackenzie, I.R., Rademakers, R., Robertson, J., Zinman, L., Rogaeva, E., 2015a. Jump from Pre-mutation to Pathologic Expansion in C9orf72. Am. J. Hum. Genet. 96, 962–970.

Xi, Z., Zhang, M., Bruni, A.C., Maletta, R.G., Colao, R., Fratta, P., Polke, J.M., Sweeney, M.G., Mudanohwo, E., Nacmias, B., Sorbi, S., Tartaglia, M.C., Rainero, I., Rubino, E., Pinessi, L., Galimberti, D., Surace, E.I., McGoldrick, P., McKeever, P., Moreno, D., Sato, C., Liang, Y., Keith, J., Zinman, L., Robertson, J., Rogaeva, E., 2015b. The C9orf72 repeat expansion itself is methylated in ALS and FTLD patients. Acta Neuropathol. 129, 715–727.

Xi, Z., Zinman, L., Moreno, D., Schymick, J., Liang, Y., Sato, C., Zheng, Y., Ghani, M., Dib, S., Keith, J., Robertson, J., Rogaeva, E., 2013. Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. Am. J. Hum. Genet. 92, 981–989.

Xiao, S., MacNair, L., McGoldrick, P., McKeever, P.M., McLean, J.R., Zhang, M., Keith, J., Zinman, L., Rogaeva, E., Robertson, J., 2015a. Isoform-specific antibodies reveal distinct subcellular localizations of C9orf72 in amyotrophic lateral sclerosis. Ann. Neurol. 78, 568–583.

Xiao, S., MacNair, L., McLean, J., McGoldrick, P., McKeever, P., Soleimani, S., Keith, J., Zinman, L., Rogaeva, E., Robertson, J., 2016. C9orf72 isoforms in Amyotrophic Lateral Sclerosis and Frontotemporal Lobar Degeneration. Brain Res.

Xiao, S., Sanelli, T., Chiang, H., Sun, Y., Chakrabartty, A., Keith, J., Rogaeva, E., Zinman, L., Robertson, J., 2015b. Low molecular weight species of TDP-43 generated by abnormal splicing form inclusions in amyotrophic lateral sclerosis and result in motor neuron death. Acta Neuropathol. 130, 49–61.

Xu, Z., Poidevin, M., Li, X., Li, Y., Shu, L., Nelson, D.L., Li, H., Hales, C.M., Gearing, M., Wingo, T.S., Jin, P., 2013. Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. Proc. Natl. Acad. Sci. 110, 7778–7783.

Yamakawa, M., Ito, D., Honda, T., Kubo, K.I., Noda, M., Nakajima, K., Suzuki, N., 2015. Characterization of the dipeptide repeat protein in the molecular pathogenesis of c9FTD/ALS. Hum. Mol. Genet. 24, 1630–1645.

Yamashita, T., Akamatsu, M., Kwak, S., 2017. Altered intracellular milieu of ADAR2-deficient motor neurons in amyotrophic lateral sclerosis. Genes (Basel).

Yang, D., Abdallah, A., Li, Z., Lu, Y., Almeida, S., Gao, F.B., 2015. FTD/ALS-associated poly(GR) protein impairs the Notch pathway and is recruited by poly(GA) into cytoplasmic inclusions. Acta Neuropathol. 130, 525–535.

Yang, M., Liang, C., Swaminathan, K., Herrlinger, S., Lai, F., Shiekhattar, R., Chen, J.-F., 2016. A C9ORF72/SMCR8-containing complex regulates ULK1 and plays a dual role in autophagy. Sci. Adv. 2, e1601167–e1601167.

Yang, W.Y., He, F., Strack, R.L., Oh, S.Y., Frazer, M., Jaffrey, S.R., Todd, P.K., Disney, M.D., 2016. Small Molecule Recognition and Tools to Study Modulation of r(CGG)exp in Fragile X-Associated Tremor Ataxia Syndrome. ACS Chem. Biol. 11, 2456–2465.

Yin, S., Lopez-Gonzalez, R., Kunz, R.C., Gangopadhyay, J., Borufka, C., Gygi, S.P., Gao, F.B., Reed, R., 2017. Evidence that C9ORF72 Dipeptide Repeat Proteins Associate with U2 snRNP to Cause Mis-splicing in ALS/FTD Patients. Cell Rep. 19, 2244–2256.

Yu, Z., Teng, X., Bonini, N.M., 2011. Triplet repeat-derived siRNAs enhance RNA-mediated toxicity in a drosophila model for myotonic dystrophy. PLoS Genet. 7, e1001340.

Zamiri, B., Reddy, K., Macgregor, R.B., Pearson, C.E., 2014. TMPyP4 porphyrin distorts RNA G-quadruplex structures of the disease-associated r(GGGGCC)n repeat of the C9orf72 gene and blocks interaction of RNAbinding proteins. J. Biol. Chem. 289, 4653–4659.

Zhang, Y.-J., Gendron, T.F., Ebbert, M.T.W., O'Raw, A.D., Yue, M., Jansen-West, K., Zhang, X., Prudencio, M., Chew, J., Cook, C.N., Daughrity, L.M., Tong, J., Song, Y., Pickles, S.R., Castanedes-Casey, M., Kurti, A., Rademakers, R., Oskarsson, B., Dickson, D.W., Hu, W., Gitler, A.D., Fryer, J.D., Petrucelli, L., 2018. Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. Nat. Med. 1.

Zhang, D., Iyer, L.M., He, F., Aravind, L., 2012. Discovery of novel DENN proteins: Implications for the evolution of eukaryotic intracellular membrane structures and human disease. Front. Genet. 3, 283.

Zhang, K., Donnelly, C.J., Haeusler, A.R., Grima, J.C., Machamer, J.B., Steinwald, P., Daley, E.L., Miller, S.J., Cunningham, K.M., Vidensky, S., Gupta, S., Thomas, M.A., Hong, I., Chiu, S.-L., Huganir, R.L., Ostrow, L.W., Matunis, M.J., Wang, J., Sattler, R., Lloyd, T.E., Rothstein, J.D., 2015. The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature 525, 56–61.

Zhang, Y.J., Jansen-West, K., Xu, Y.F., Gendron, T.F., Bieniek, K.F., Lin, W.L., Sasaguri, H., Caulfield, T., Hubbard, J., Daughrity, L., Chew, J., Belzil, V. V., Prudencio, M., Stankowski, J.N., Castanedes-Casey, M., Whitelaw, E., Ash, P.E.A., DeTure, M., Rademakers, R., Boylan, K.B., Dickson, D.W., Petrucelli, L., 2014. Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. Acta Neuropathol. 128, 505–524.

Zhang, Y.-J., Gendron, T.F., Grima, J.C., Sasaguri, H., Jansen-West, K., Xu, Y.-F., Katzman, R.B., Gass, J., Murray, M.E., Shinohara, M., Lin, W.-L., Garrett, A., Stankowski, J.N., Daughrity, L., Tong, J., Perkerson, E.A., Yue, M., Chew, J., Castanedes-Casey, M., Kurti, A., Wang, Z.S., Liesinger, A.M., Baker, J.D., Jiang, J., Lagier-Tourenne, C., Edbauer, D., Cleveland, D.W., Rademakers, R., Boylan, K.B., Bu, G., Link, C.D., Dickey, C.A., Rothstein, J.D., Dickson, D.W., Fryer, J.D., Petrucelli, L., 2016. C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. Nat. Neurosci. 19, 668–677.

Zhou, B., Liu, C., Geng, Y., Zhu, G., 2015. Topology of a G-quadruplex DNA formed by C9orf72 hexanucleotide repeats associated with ALS and FTD. Sci. Rep. 5, 16673.

Zhu, J., Cynader, M.S., Jia, W., 2015. TDP-43 inhibits NF-κB activity by blocking p65 nuclear translocation. PLoS One 10, e0142296.

Zu, T., Cleary, J.D., Liu, Y., Bañez-Coronel, M., Bubenik, J.L., Ayhan, F., Ashizawa, T., Xia, G., Clark, H.B., Yachnis, A.T., Swanson, M.S., Ranum, L.P.W., 2017. RAN Translation Regulated by Muscleblind Proteins in Myotonic Dystrophy Type 2. Neuron 95, 1292–1305.e5.

Zu, T., Gibbens, B., Doty, N.S., Gomes-pereira, M., Huguet, A., Stone, M.D., 2010. Non-ATG – initiated translation directed by microsatellite expansions. Pnas 108, 260–265.

Zu, T., Liu, Y., Banez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., Miller, T.M., Harms, M.B., Falchook, A.E., Subramony, S.H., Ostrow, L.W., Rothstein, J.D., Troncoso, J.C., Ranum, L.P.W., 2013. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc. Natl. Acad. Sci. 110, E4968–E4977.

Appendix

Appendix 1: Quantification of DsRed2 protein levels

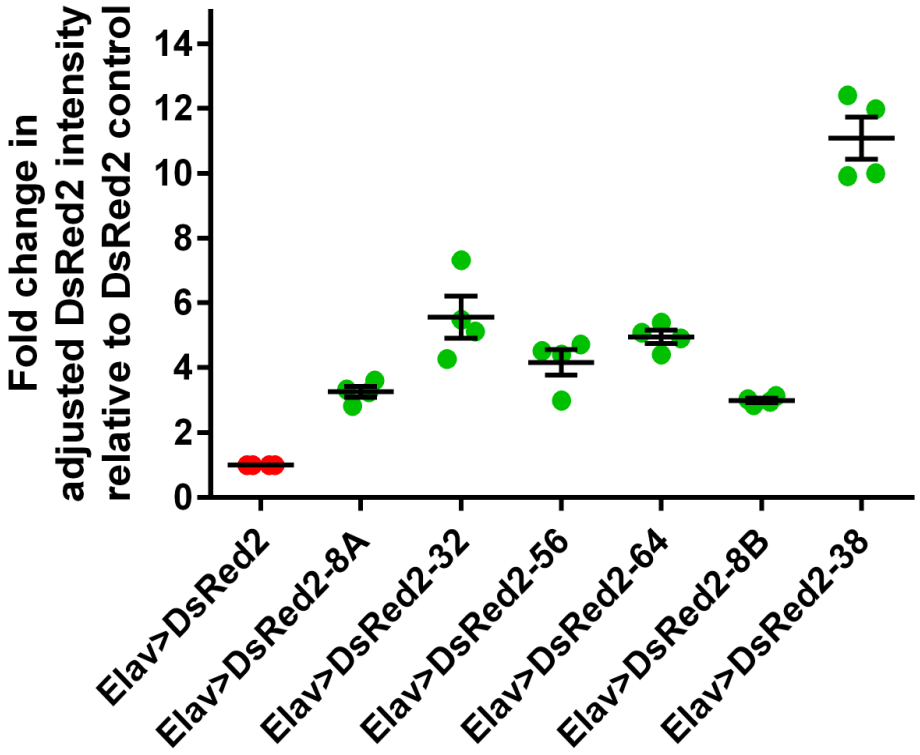

Figure A1| The presence of G_4C_2 repeats in the 3'UTR of DsRed2 correlates with increased DsRed2 expression levels. Western blot analysis of RIPA soluble protein from day 5 head extracts from flies expressing DsRed2. A one-way ANOVA was used to compare DsRed2 expression levels. Levels of DsRed2 were significantly affected by genotype ($F=49.88$, $p<0.0001$). Bonferroni-Holm corrected multiple comparisons can be found in appendix 7, table A7. This revealed flies expressing DsRed2-38 repeats had significantly higher levels of DsRed2 compared to all over genotypes. There were no significant differences in DsRed2 expression levels between the 32 vs 56, 32 vs 64 and 56 vs 64 repeat lines. Intensity of DsRed2 signal was adjusted with respect to the signal intensity from the corresponding tubulin loading control. Mean with SEM is shown for each genotype (n=4). This figure was generated by Alan Stepto and used with his permission.

Appendix 2: Toxicity in the 38 repeat line is not attributable to increased DsRed2 expression levels

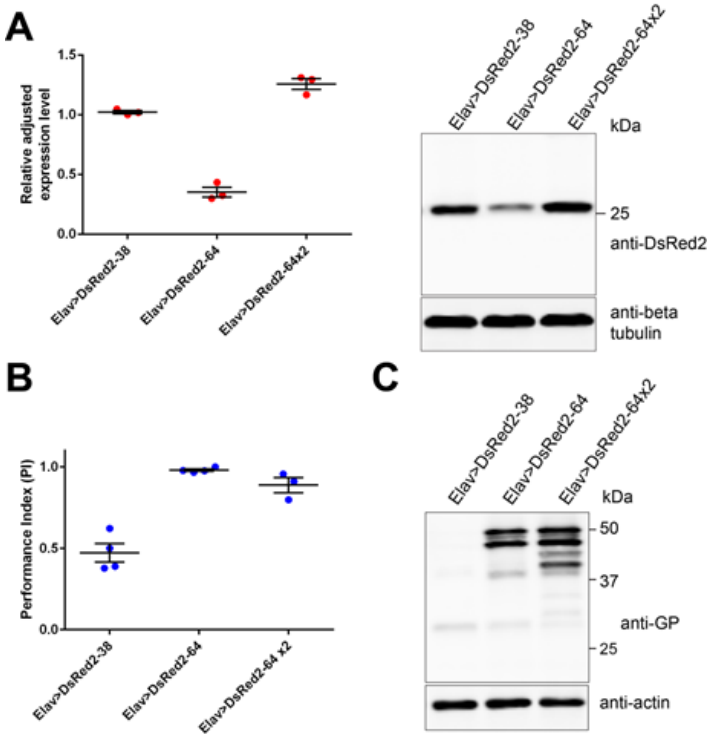

Figure A2 | Increased G_4C_2 transcript and DsRed2 protein levels do not explain neurotoxicity in the 38 G_4C_2 repeat line (A) Western blot analysis of RIPA soluble protein from day 5 head extracts from flies expressing DsRed2-38 repeats (*Elav>DsRed2(G_4C_2)38*), one copy of the DsRed2-64 repeats (*Elav>DsRed2(G_4C_2)64*) and two copies of the DsRed2-64 repeats (Elav>DsRed2(G_4C_2)64x2). A one-way ANOVA revealed DsRed2 expression level was significantly affected by genotype ($F=168.7$, $p<0.0001$). Bonferroni-Holm corrected multiple comparisons can be found in appendix 7, table A7. This revealed flies expressing two copies of DsRed2-64 produced significantly higher levels of DsRed2 than flies harbouring one copy of the DsRed2-64 repeat construct and also flies expressing DsRed2-38 repeats. Mean with SEM is shown for each genotype (n=3). **(B)** A one-way ANOVA was used to compare the mean differences in climbing performance (the performance index - PI) between flies expressing DsRed2-38 repeats, one copy of the DsRed2-64 repeats and two copies of the DsRed2-64 repeats at day 5. The PI of flies was significantly affected by both genotype ($F=43.28$, $p<0.0001$). Bonferroni-Holm corrected multiple comparisons can be found in appendix 7, table A7. This revealed flies expressing DsRed2-38 repeats had a significantly impaired climbing performance compared to flies with one copy of DsRed2-64 repeats and two copies of DsRed2-64 repeats. Each data point represents the PI of ~15 flies. Mean with SEM is shown for each genotype (n=3). **(C)** Two copies of the DsRed2-64 construct increases the number of poly-GP bands detected compared to flies with one copy of the DsRed2-64 construct. These experiments and this figure were generated by Alan Stepto and used with his permission.

Appendix 3: Validation of a novel monoclonal poly-GP antibody

To investigate RAN translation in the hexanucleotide repeat expansion *Drosophila* model used in this study, a monoclonal antibody for the DPR poly-GP was generated. This monoclonal poly-GP antibody was generated by the immunisation of mice with the following peptide sequence - GPGPGPGPGPGPGPGPGPGP (GPx10) by Abmart, China. Poly-GP is highly expressed in the C9ALS/FTD CNS with poly-GP inclusions being the second most abundant DPR aggregate found in the post-mortem brain after poly-GA (Schludi et al., 2015) and has strong potential to be a pharmacodynamic marker for C9ALS/FTD due to its solubility (Gendron et al., 2015; Gendron et al., 2017; Lehmer et al., 2017). Hence the generation of a highly specific monoclonal antibody for poly-GP is not only an important tool for investigating RAN translation *in vivo* and DPR pathology in post-mortem tissue, but also has important clinical implications.

To confirm the specificity of the novel monoclonal poly-GP antibody produced by Abmart, lysate from HEK293 cells expressing either - poly-GP, poly-GA, poly-GR, poly-PA or poly-PR from alternative non G_4C_2 codons (provided by Younbok Lee) was probed with the monoclonal poly-GP antibody. Dot blotting showed a specific high intensity signal for the poly-GP HEK293 cell lysate only (figure 3.4A3 A). No signal was seen for the poly-GA, poly-GR, poly-PA or poly-PR lysate. Following confirmation that the antibody was specific for the poly-GP peptide and does not cross-react with any of the other DPRs, immunofluorescent staining of human cerebellum post-mortem tissue was carried out to determine if the antibody could detect inclusions characteristic of the C9ALS/FTD brain (figure 3.4A3 B). Staining with the monoclonal poly-GP antibody detected the presence of neuronal cytoplasmic inclusions of poly-GP in the C9ALS/FTD cerebellum, particularly in the granule layer as has been previously reported (Schludi et al., 2015). No poly-GP inclusions were found in Alzheimer's disease (AD) cerebellum. Hence this novel monoclonal poly-GP antibody can detect poly-GP with high affinity and specificity and was therefore suitable for investigating RAN translation in the brain of flies harbouring the repeat expansion.

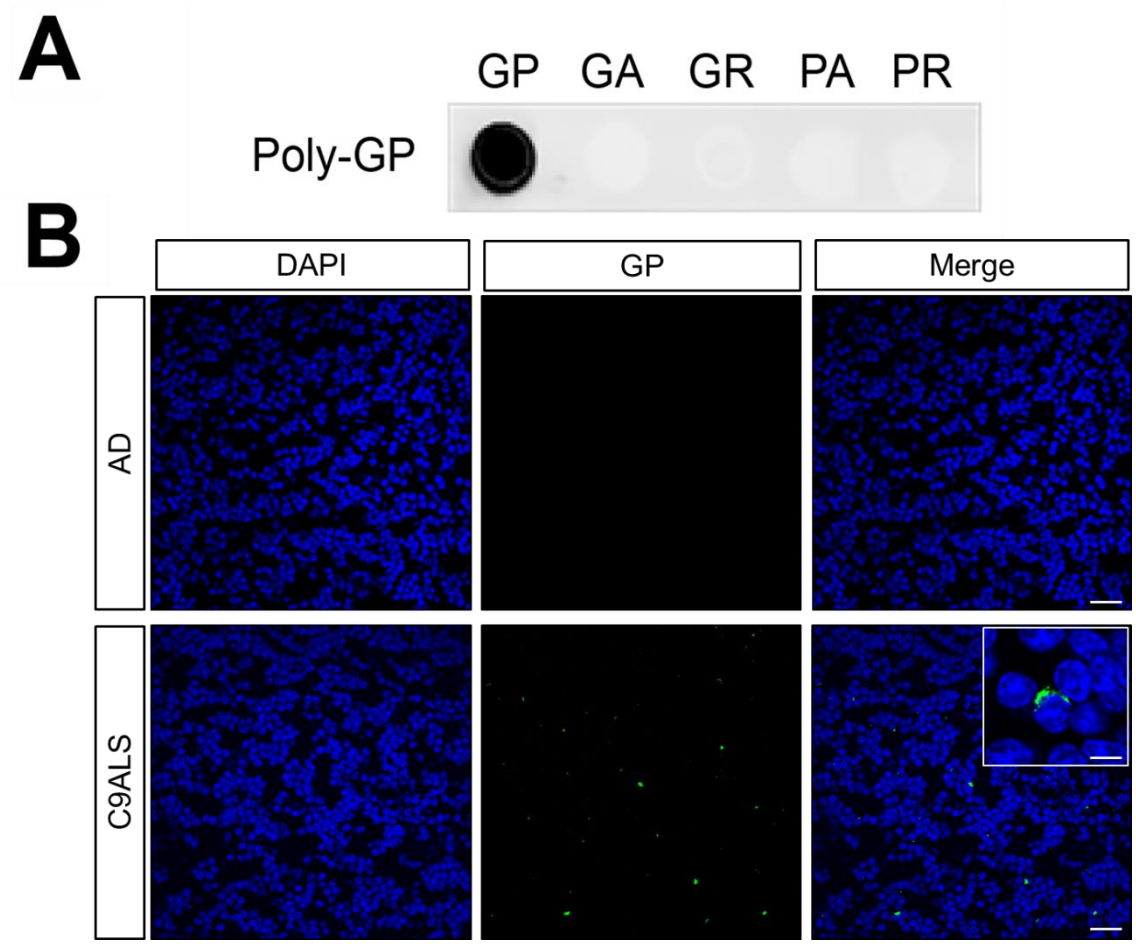

Figure A3 | Validation of novel monoclonal poly-GP antibody. (A) Monoclonal poly-GP antibody specifically detects lysate from HEK293 cells expressing GP, but not poly-GA, poly-GR, poly-PA or poly-PR. **(B)** Staining of Alzheimer's disease (AD) and C9ALS/FTD cerebellum revealed the presence of poly-GP inclusions in the granule layer of the cerebellum but not in AD cerebellum. Scale bars, 25 µm; Scale bar for inset, 5 µm.

Appendix 4: The expression pattern and levels of of poly-GP does not change over time

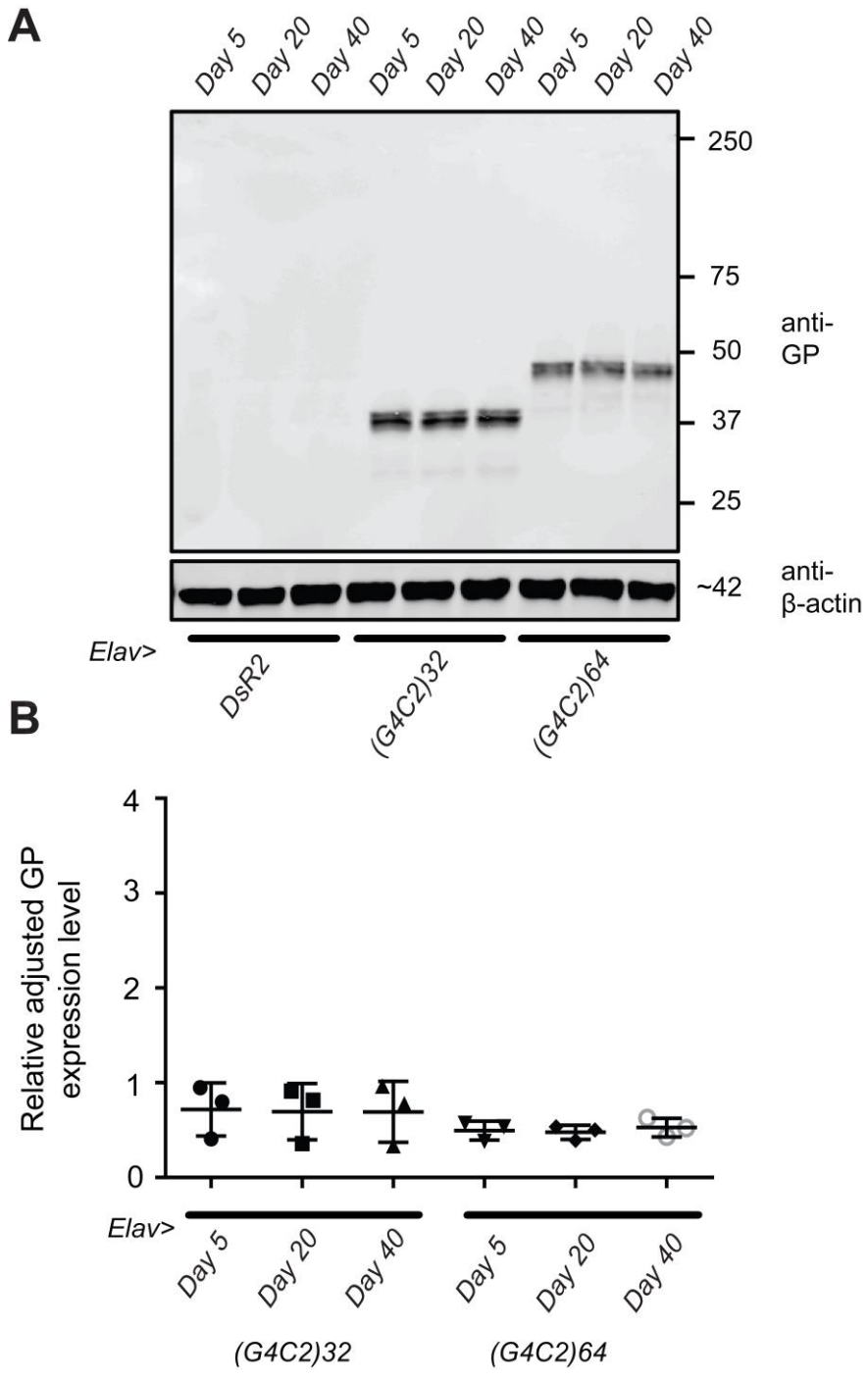

Figure A4 | Poly-GP levels do not change over time. **(A)** Western blot analysis of RIPA soluble protein from day 5, 20 and 40 head extracts from flies expressing DsRed2 (*Elav>DsR2*), 38 G_4C_2 repeats (Elav>DsRed2(G_4C_2)38), 64 G_4C_2 repeats (Elav>DsRed2(G_4C_2)64). **(B)** Quantification of western blots for expression of poly-GP. A one-way ANOVA demonstrated there to be no significant difference at the different time points for poly-GP expression levels in the 32 ($F=0.006913$, $P=0.9931$) and 64 ($F=0.2274$, $P=0.8032$) G_4C_2 repeat lines. Mean with SEM is shown for each time point (n=3).

Appendix 5: Full length sequences of the DPR constructs generated

GA8

ATGTACCCATACGACGTCCCAGACTACGCTGGAGCTGGTGCTGGAGCAGGTGCTGGAGCAGGTGCTGGAGCAGGAGCTG
AGCAAAAGCTTATTTCTGAAGAGGACTTGTAA

GA64

ATGTACCCATACGACGTCCCAGACTACGCTGGAGCTGGTGCTGGAGCAGGTGCTGGAGCAGGTGCTGGAGCAGGAGCTG
GTGCAGGTGCTGGAGCAGGAGCTGGTGCTGGAGCAGGAGCAGGTGCTGGAGCTGGTGCAGGAGCAGGTGCTGGAGCAG
GTGCTGGAGCAGGTGCTGGAGCTGGTGCAGGAGCAGGAGCAGGTGCTGGTGCTGGTGCAGGAGCTGGTGCAGGAGCTG
GAGCAGGTGCAGGAGCTGGTGCTGGAGCAGGTGCAGGTGCTGGAGCAGGAGCAGGTGCAGGTGCTGGAGCTGGTGCAG
GAGCTGGTGCAGGAGCTGGTGCAGGTGCTGGTGCAGGAGCTGGTGCAGGAGCTGGTGCAGGAGCTGGTGCAGGTGCAG
GAGCTGGTGCTGGAGCAGGAGCTGAGCAAAAGCTTATTTCTGAAGAGGACTTGTAA

GR8

ATGTACCCATACGACGTCCCAGACTACGCTGGTCGTGGAAGAGGTCGAGGTAGAGGTCGAGGAAGAGGACGTGGTCGTG
AGCAAAAGCTTATTTCTGAAGAGGACTTG

GR64

ATGTACCCATACGACGTCCCAGACTACGCTGGTCGTGGAAGAGGTCGAGGTAGAGGTCGAGGAAGAGGACGTGGTCGTG
GTAGAGGACGAGGTAGAGGACGAGGACGTGGTAGAGGAAGAGGACGAGGACGTGGTAGAGGTAGAGGACGTGGAAGAG
GTCGAGGACGTGGTAGAGGAAGAGGTCGTGGACGTGGTAGAGGTCGAGGTCGTGGACGAGGTAGAGGTAGAGGACGTG
GTAGAGGACGAGGAAGAGGTCGTGGTCGAGGACGAGGACGTGGTAGAGGTAGAGGACGAGGAAGAGGTCGAGGTAGAG
GACGTGGACGAGGACGTGGTAGAGGAAGAGGACGAGGTCGAGGTAGAGGAAGAGGACGAGGTCGTGGTAGAGGACGTG
GAAGAGGTCGAGGTCGAGGAAGAGAGCAAAAGCTTATTTCTGAAGAGGACTTG

PR8

ATGTACCCATACGACGTCCCAGACTACGCTCCTGCTCCAGCTCCTGCACCTGCACCAGCACCTGCTCCAGCTCCTGCAGA
GCAAAAGCTTATTTCTGAAGAGGACTTG

PA64

ATGTACCCATACGACGTCCCAGACTACGCTCCTGCTCCAGCTCCTGCACCTGCACCAGCACCTGCTCCAGCTCCTGCACC
AGCTCCTGCACCAGCACCAGCACCAGCTCCAGCACCTGCACCAGCTCCTGCTCCTGCACCAGCTCCTGCTCCTGCACCAG
CTCCTGCTCCTGCACCAGCACCAGCACCTGCTCCTGCACCTGCACCTGCTCCAGCACCAGCTCCTGCTCCAGCACCAGCA
CCTGCACCAGCTCCTGCACCTGCTCCAGCTCCTGCACCAGCTCCAGCTCCAGCACCTGCTCCAGCACCTGCACCTGCACC
AGCTCCAGCTCCTGCTCCAGCACCAGCTCCAGCTCCTGCTCCTGCACCAGCTCCTGCACCTGCTCCTGCACCAGCACCAG
CTCCTGCACCAGCTGAGCAAAAGCTTATTTCTGAAGAGGACTTG

PR8

ATGTACCCATACGACGTCCCAGACTACGCTCCTAGACCTCGTCCAAGGCCAAGACCTCGTCCACGACCAAGGCCTAGAGA
GCAAAAGCTTATTTCTGAAGAGGACTTG

PR64

ATGTACCCATACGACGTCCCAGACTACGCTCCTAGACCTCGTCCAAGGCCAAGACCTCGTCCACGACCAAGGCCTAGACC
TCGTCCTAGACCTAGGCCAAGGCCTAGACCTCGACCAAGGCCTCGTCCACGACCTAGGCCAAGACCTCGTCCACGACCTA
GGCCTAGACCACGTCCTCGACCAAGGCCAAGACCTAGACCTCGTCCACGACCTAGACCAAGGCCTCGACCTAGGCCACG
TCCTCGACCAAGACCAAGGCCACGACCTAGACCTCGTCCACGTCCTAGACCTCGTCCTAGGCCAAGACCTCGACCAAGGC
CAAGACCTCGTCCAAGACCTAGGCCACGACCTAGACCAAGGCCTCGACCTCGTCCAAGACCACGTCCTAGACCAAGGCCT
CGTCCACGTCCAAGAGAGCAAAAGCTTATTTCTGAAGAGGACTTG

Appendix 6: *TBPH* localisation in pupal eye discs expressing poly-GR64

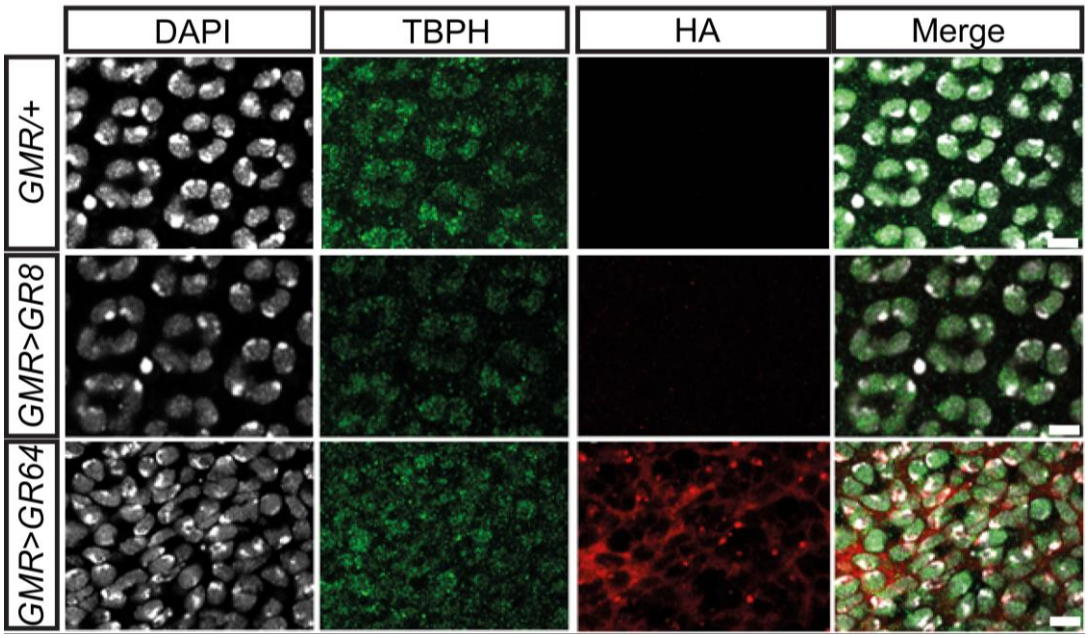

Figure A6 | Cytosolic localisation of *TBPH* is increased in the pupal eye disc of poly-GR expressing flies. Pupal eye discs were immunolabeled with *TBPH*. No changes were seen in localisation of *TBPH* in controls (*GMR/+* and *GMR>GR8*). TBPH was seen in accumulate in the cytoplasm of poly-GR64 expressing (*GMR>GR64*) pupal eye discs. Scale bars, 10µm.

Appendix 7: ΔNLS-*TBPH* expression causes a cytoplasmic accumulation and concomitant nuclear reduction of endogenous *TBPH*

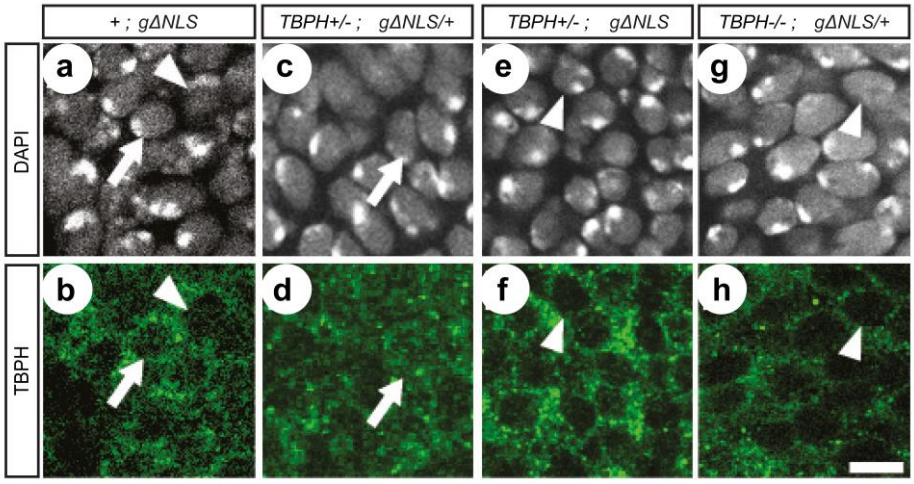

Figure A7 | ΔNLS-*TBPH* (gΔNLS) expression causes a cytoplasmic accumulation of *TBPH* and a concomitant nuclear reduction. (a-h) Confocal images of third instar larval eye discs immunolabeled with anti-*TBPH*. Flies homozygous gΔNLS and those heterozygous for a deletion of *TBPH* that also express one copy of a genomic (gΔNLS) construct gΔNLS reveal **(a, b, c, d, arrow)** diffuse cytoplasmic and nuclear *TBPH* localisation; in heterozygous *TBPH* mutants with two copies of ΔNLS-*TBPH*, nuclear clearance of *TBPH* is much more pronounced **(e, f, arrowhead)** and is also seen **(g, h)** in homozygous *TBPH* null mutants expressing one copy of ΔNLS-*TBPH* **(i).** Scale bars, 10µm. Experiments performed by Dr. Dani Diaper.

Appendix 8: Poly-GP detectable in axons at the larval neuromuscular junction

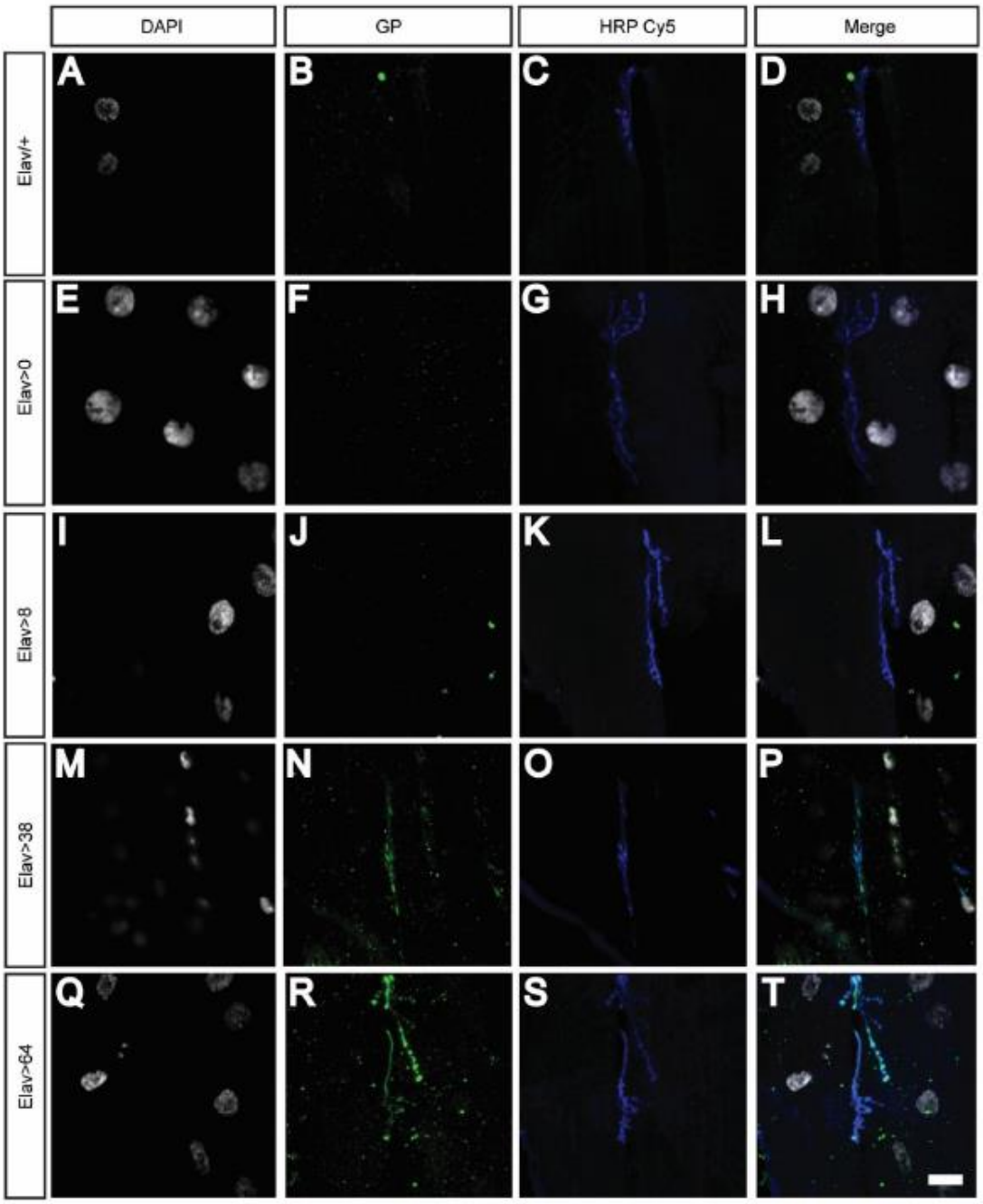

Figure A8 | Poly-GP is detectable in motor neuron axons of the larval neuromuscular junction. Larval neuromuscular junctions were dissected and stained for poly-GP and HRP Cy5 which detects motor neurons. No poly-GP signal was seen in the motor neuron axons of control genotypes (*Elav/+*, *Elav>0* and *Elav>8*). Poly-GP was however detectable in the axons of motor neurons of flies expressing 38 repeats (*Elav>38*) and 64 repeats (*Elav>64*). Scale bars, 10µm. These experiments and this figure were generated by Wing Hei Au and used with her permission.

Appendix 9: Post mortem cases used in this study

Table A9.1 | Details of the post-mortem cases used in this study

Case Number	MRC ID	Diagnosis	Age	Sex	PMD
A145/02	BBN_15765	Control	55	M	24
A135/95	BBN_16621	Control	65	M	24
A154/94	BBN_16733	Control	55	F	24
A278/96	BBN_16525	Control	77	F	29
A205/94	BBN_16707	Control	71	F	30
A281/00	BBN_15274	FTD C9-VE	61	F	5
A163/07	BBN_15288	FTD C9-VE	80	M	6
A103/08	BBN_15294	FTD C9-VE	85	F	24
A370/97	BBN_16467	FTD C9-VE	57	M	2
A265/00	BBN_15273	FTD C9-VE	50	F	35
A315/07	BBN_15291	FTD C9+VE	56	M	19
A163/09	BBN_15300	FTD C9+VE	79	M	35
A088/00	BBN_15270	FTD C9+VE	57	F	16
A112/99	BBN_16438	FTD C9+VE	58	F	12
A117/91	BBN_16969	FTD C9+VE	57	F	12
A144/12	BBN_4253	ALS/FTD C9+VE	59	F	21
A205/07	BBN_9863	AD Braak VI	M	73	24

*Green used for western blotting only *Blue used for IHC only *Red used for IHC and western blotting

Table A9.2 | Details of the post-mortem cases used in this study

Case Number	Diagnosis	% of neuronal staining	% of nuclear staining	% of cytoplasmic staining	Lack of nuclear staining	Dystrophic neurites	Inclusions	Background staining	Comments
A163/07	FTD C9-VE	60	20	40	Mild	-	-	1	
A103/08	FTD C9-VE	50	20	30	Mild	-	-	0	
A315/07	FTD C9+VE	70	20	50	Marked	+Thin and elongated	-	1	
A163/09	FTD C9+VE	60	40	20	Moderate	+Thin and elongated	+ Occa-sional	1	Cytoplasmic positivity diffuse and granular

* Tissue analysis performed by neuropathologist Dr. Tibor Hortobagyi includes: % of neuronal staining, % of nuclear staining, % of cytoplasmic staining, lack of nuclear staining (marked, mild, moderate), dystrophic neurites (+/- and description), inclusions (+/-) and background staining (0–none, 1-mild, 2-moderate, 3-intense).

Appendix 10: Pairwise comparisons

Table A10 | Bonferroni-Holm adjusted pairwise comparisons

Figure	Test used	Comparison	Non-adjusted p-value	Bonferroni-Holm adjusted pairwise comparison
A1	One-way ANOVA	Elav>DsRed2-8A vs Elav>DsRed2-38	<0.0001	0.003333333
		Elav>DsRed2-32 vs Elav>DsRed2-38	<0.0001	0.003571429
		Elav>DsRed2-56 vs Elav>DsRed2-38	<0.0001	0.003846154
		Elav>DsRed2-64 vs Elav>DsRed2-38	<0.0001	0.004166667
		Elav>DsRed2-8B vs Elav>DsRed2-38	<0.0001	0.004545455
		Elav>DsRed2-32 vs Elav>DsRed2-8B	0.0004	0.005
		Elav>DsRed2-8A vs Elav>DsRed2-32	0.0012	0.005555556
		Elav>DsRed2-64 vs Elav>DsRed2-8B	0.0042	0.00625

		Elav>DsRed2-8A vs Elav>DsRed2-64	0.0109	0.007142857
		Elav>DsRed2-32 vs Elav>DsRed2-56	0.0313	0.008333333
		Elav>DsRed2-56 vs Elav>DsRed-8B	0.654	0.01
		Elav>DsRed2-8A vs Elav>DsRed2-56	0.1452	0.0125
		Elav>DsRed2-56 vs Elav>DsRed2-64	0.2052	0.016666667
		Elav>DsRed2-32 vs Elav>DsRed2-64	0.3208	0.025
		Elav>DsRed-8A vs Elav>DsRed-8B	0.6658	0.05
3.5B	One-way ANOVA	Elav>DsRed2-56 vs Elav>DsRed2-64	<0.0001	0.0083333
		Elav>DsRed2-64 vs Elav>DsRed2-38	<0.0001	0.01
		Elav>DsRed2-32 vs Elav>DsRed2-56	<0.0001	0.0125
		Elav>DsRed2-32 vs Elav>DsRed2-64	0.0001	0.0166667
		Elav>DsRed2-32 vs Elav>DsRed2-38	0.0037	0.025
		Elav>DsRed2-56 vs Elav>DsRed2-38	0.0063	0.05

3.6B	One-way ANOVA	Elav>DsRed2-64 vs Elav>DsRed2-38	<0.0001	0.0083333
		Elav>DsRed2-32 vs Elav>DsRed2-64	<0.0001	0.01
		Elav>DsRed2-32 vs Elav>DsRed2-38	<0.0001	0.0125
		Elav>DsRed2-56 vs Elav>DsRed2-64	0.0001	0.0166667
		Elav>DsRed2-56 vs Elav>DsRed2-38	0.0052	0.025
		Elav>DsRed2-32 vs Elav>DsRed2-56	0.0062	0.05
4.3A & C	Two-way ANOVA	Elav/+ vs Elav>DsRed2>38	<0.0001	0.001785714
		Elav>DsRed2 vs Elav>DsRed2-38	<0.0001	0.001851852
		Elav>DsRed2-8A vs Elav>DsRed2-38	<0.0001	0.001923077
		Elav>DsRed2-8B vs Elav>DsRed2-38	<0.0001	0.002
		Elav>DsRed2-32 vs Elav>DsRed2-38	<0.0001	0.002083333

	Elav>DsRed2-56 vs Elav>DsRed2-38	<0.0001	0.002173913
	Elav>DsRed2-64 vs Elav>DsRed2-38	<0.0001	0.002272727
	Elav/+ vs Elav>DsRed2>8A	0.0883	0.002380952
	Elav>DsRed2-8A vs Elav>DsRed2-8B	0.1495	0.0025
	Elav/+ vs Elav>DsRed2>32	0.2382	0.002631579
	Elav/+ vs Elav>DsRed2>56	0.2764	0.002777778
	Elav>DsRed2 vs Elav>DsRed2-8A	0.3192	0.002941176
	Elav>DsRed2-8B vs Elav>DsRed2-32	0.3626	0.003125
	Elav>DsRed2-8A vs Elav>DsRed2-64	0.3648	0.003333333
	Elav/+ vs Elav>DsRed2-64	0.413	0.003571429

Elav>DsRed2-8B vs Elav>DsRed2-56	0.413	0.003846154
Elav/+ vs Elav>DsRed2	0.4664	0.004166667
Elav>DsRed2-8A vs Elav>DsRed2-56	0.5254	0.004545455
Elav>DsRed2-8B vs Elav>DsRed2-64	0.5844	0.005
Elav>DsRed2-8A vs Elav>DsRed2-32	0.5874	0.005555556
Elav>DsRed2 vs Elav>DsRed2-32	0.6472	0.00625
Elav>DsRed2 vs Elav>DsRed2-8B	0.6484	0.007142857
Elav>DsRed2-32 vs Elav>DsRed2-64	0.7139	0.008333333
Elav>DsRed2 vs Elav>DsRed2-56	0.7151	0.01

		Elav/+ vs Elav>DsRed2-8B	0.7842	0.0125
		Elav>DsRed2-56 vs Elav>DsRed2-64	0.7842	0.016666667
		Elav>DsRed2-32 vs Elav>DsRed2-56	0.9259	0.025
		Elav>DsRed2 vs Elav>DsRed2-64	0.9273	0.05
4.1B & D	Two-way ANOVA	Elav/+ vs Elav>DsRed2>38	<0.0001	0.001785714
		Elav>DsRed2 vs Elav>DsRed2-38	<0.0001	0.001851852
		Elav>DsRed2-8A vs Elav>DsRed2-38	<0.0001	0.001923077
		Elav>DsRed2-8B vs Elav>DsRed2-38	<0.0001	0.002
		Elav>DsRed2-32 vs Elav>DsRed2-38	<0.0001	0.002083333
		Elav>DsRed2-56 vs Elav>DsRed2-38	<0.0001	0.002173913

Elav>DsRed2-64 vs Elav>DsRed2-38	<0.0001	0.002272727
Elav/+ vs Elav>DsRed2>56	0.016	0.002380952
Elav>DsRed2-8A vs Elav>DsRed2-56	0.0416	0.0025
Elav/+ vs Elav>DsRed2>64	0.0541	0.002631579
Elav>DsRed2 vs Elav>DsRed2-56	0.0626	0.002777778
Elav/+ vs Elav>DsRed2-8B	0.0627	0.002941176
Elav>DsRed2-32 vs Elav>DsRed2-56	0.1017	0.003125
Elav>DsRed2-8A vs Elav>DsRed2-64	0.15	0.003333333
Elav>DsRed2-8A vs Elav>DsRed2-8B	0.1683	0.003571429
Elav>DsRed2 vs Elav>DsRed2-64	0.207	0.003846154

Elav>DsRed2 vs Elav>DsRed2-8B	0.2212	0.004166667
Elav/+ vs Elav>DsRed2>32	0.3082	0.004545455
Elav>DsRed2-32 vs Elav>DsRed2-64	0.3203	0.005
Elav>DsRed2-8B vs Elav>DsRed2-32	0.3318	0.005555556
Elav/+ vs Elav>DsRed2>8A	0.426	0.00625
Elav>DsRed2-56 vs Elav>DsRed2-64	0.4275	0.007142857
Elav/+ vs Elav>DsRed2	0.4447	0.008333333
Elav>DsRed2-8B vs Elav>DsRed2-56	0.4654	0.01
Elav>DsRed2-8A vs Elav>DsRed2-32	0.7282	0.0125

		Elav>DsRed2 vs Elav>DsRed2-32	0.7844	0.016666667
		Elav>DsRed2 vs Elav>DsRed2-8A	0.9648	0.025
		Elav>DsRed2-8B vs Elav>DsRed2-64	0.9734	0.05
A2 A	One-way ANOVA	Elav>DsRed2-64 vs Elav>DsRed2-64x2	<0.0001	0.0166667
		Elav>DsRed2-38 vs Elav>DsRed2-64	<0.0001	0.025
		Elav>DsRed2-38 vs Elav>DsRed2-64x2	0.0037	0.05
A2 B	One-way ANOVA	Elav>DsRed2-64 vs Elav>DsRed2-38	<0.0001	0.0166667
		Elav>DsRed2-64x2 vs Elav>DsRed2-38	0.0002	0.025
		Elav>DsRed2-64 vs Elav>DsRed2-64x2	0.1756	0.05
4.1E	One-way ANOVA	Elav/+ vs Elav>DsRed2>64	<0.0001	0.003333333
		Elav>DsRed2 vs Elav>DsRed2-32	<0.0001	0.003571429
		Elav>DsRed2 vs Elav>DsRed2-64	<0.0001	0.003846154

		Elav>DsRed2-56 vs Elav>DsRed2-64	0.0003	0.004166667
		Elav/+ vs Elav>DsRed2>32	0.0005	0.004545455
		Elav>DsRed2-56 vs Elav>DsRed2-32	0.0014	0.005
		Elav>DsRed2-8A vs Elav>DsRed2-64	0.0023	0.005555556
		Elav>DsRed2-8A vs Elav>DsRed2-32	0.0068	0.00625
		Elav>DsRed2 vs Elav>DsRed2-8A	0.0469	0.007142857
		Elav>DsRed2 vs Elav>DsRed2-56	0.1742	0.008333333
		Elav/+ vs Elav>DsRed2-8A	0.29	0.01
		Elav/+ vs Elav>DsRed2	0.3612	0.0125
		Elav>DsRed2-8A vs Elav>DsRed2-56	0.5177	0.016666667
		Elav/+ vs Elav>DsRed2-56	0.67	0.025
		Elav>DsRed2-32 vs Elav>DsRed2-64	0.8354	0.05
4.2B	One-way ANOVA	Elav/+ vs Elav>DsRed2-38	<0.0001	0.0166667

		Elav>DsRed2-8B vs Elav>DsRed2-38	<0.0001	0.025
		Elav/+ vs Elav>DsRed2-8B	0.5006	0.05
4.3B	One-way ANOVA	Elav>DsRed2 vs Elav>DsRed2-32	<0.0001	0.005
		Elav>DsRed2 vs Elav>DsRed2-64	<0.0001	0.003571429
		Elav>DsRed2-8A vs Elav>DsRed2-32	0.0001	0.00625
		Elav>DsRed2-8A vs Elav>DsRed2-64	0.0006	0.007142857
		Elav>DsRed2-32 vs Elav>DsRed2-56	0.0015	0.008333333
		Elav>DsRed2-56 vs Elav>DsRed2-64	0.0062	0.01
		Elav>DsRed2 vs Elav>DsRed2-56	0.0515	0.0125
		Elav>DsRed2-8A vs Elav>DsRed2-56	0.3509	0.016666667
		Elav>DsRed2-32 vs Elav>DsRed2-64	0.4006	0.025
		Elav>DsRed2 vs Elav>DsRed2-8A	0.4093	0.05
4.4B	Chi-square	Elav>DsRed2 vs Elav>DsRed2-8B	0.7701	n/a

		Elav>DsRed2 vs Elav>DsRed2-64	0.8816	n/a
		Elav>DsRed2 vs Elav>DsRed2-38	<0.0001	n/a
		Elav>DsRed2-8B vs Elav>DsRed2-64	0.9924	n/a
		Elav>DsRed2-8B vs Elav>DsRed2-38	<0.0001	n/a
		Elav>DsRed2-64 vs Elav>DsRed2-38	<0.0001	n/a
4.4C	One-way ANOVA	Elav>DsRed2-8B vs Elav>DsRed2-38	<0.0001	0.008333333
		Elav>DsRed2-64 vs Elav>DsRed2-38	<0.0001	0.01
		Elav>DsRed2 vs Elav>DsRed2-38	<0.0001	0.0125
		Elav>DsRed2 vs Elav>DsRed2-8B	0.4188	0.016666667
		Elav>DsRed2 vs Elav>DsRed2-64	0.4650	0.025
		Elav>DsRed2-8B vs Elav>DsRed2-64	0.9623	0.05
6.2B	One-way ANOVA	Elav>NLS/+ vs Elav>NLS/DsRed2-38	<0.0001	0.003333333
		Elav>DsRed2-64 vs Elav>NLS/DsRed2-38	<0.0001	0.003571429

Elav>/+ vs Elav>NLS/DsRed2-38	<0.0001	0.003846154
Elav>NLS/DsRed2 vs Elav>NLS/DsRed2-38	<0.0001	0.004166667
Elav>NLS/+ vs Elav>NLS/DsRed2-38	<0.0001	0.004545455
Elav>DsRed2-64 vs Elav>DsRed2-38	<0.0001	0.005
Elav/+ vs Elav>DsRed2-38	<0.0001	0.005555556
Elav>NLS/DsRed2-64 vs Elav>DsRed2-38	<0.0001	0.00625
Elav>DsRed2-38 vs Elav>NLS/DsRed2-38	<0.0001	0.007142857
Elav>NLS/+ vs Elav>NLS-DsRed2-64	0.0946	0.008333333
Elav>DsRed-64 vs Elav>NLS/DsRed2-64	0.1161	0.01
Elav/+ vs Elav>NLS/DsRed-64	0.3181	0.0125
Elav/+ vs Elav>NLS/+	0.4702	0.016666667
Elav/+ vs Elav>DsRed-64	0.5404	0.025
Elav>NLS/+ vs Elav>DsRed2	0.9108	0.05

6.3B	Unpaired t-test	Elav>DsRed-64 vs Elav>NLS/DsRed2-64	0.0004	n/a
		Elav>DsRed2-38 vs Elav>NLS/DsRed2-38	0.0555	n/a
6.3D	Unpaired t-test	Elav>DsRed-64 vs Elav>NLS/DsRed2-64	0.0003	n/a
		Elav>DsRed2-38 vs Elav>NLS/DsRed2-38	0.0098	n/a